STUDENTS, TEACHERS, AND A FUNDAMENTAL MISMATCH: COMPARING PERSPECTIVES ON THE PURPOSES OF PUNISHMENT IN SCHOOL SETTINGS

Vom Promotionsausschuss des Fachbereichs 8: Psychologie der Universität Koblenz-Landau

zur Verleihung des akademischen Grades Doktor der Philosophie (Dr. phil.)

genehmigte Dissertation

vorgelegt von:

Mathias Twardawski, M. Sc.

Vorsitzende des Promotionsausschusses:

Prof. Dr. Melanie Steffens (Universität Koblenz-Landau)

Vorsitzende der Promotionskommission:

Jun. Prof. Dr. Meike Kroneisen (Universität Koblenz-Landau)

Berichterstatterin und Berichterstatter:

Prof. Dr. Anna Baumert (Max Planck Institut zur Erforschung von Gemeinschaftsgütern)

Prof. Dr. Manfred Schmitt (Universität Koblenz-Landau)

Datum der wissenschaftlichen Aussprache: 08. März 2019

“It is to Madame Justice that I dedicate this concerto,

in honor of the holiday that she seems to have taken from these parts,

and in recognition of the impostor that stands in her stead.”

V

- V for Vendetta -

CONTENT

List of tables 5

List of figures 6

Abstract 7

Zusammenfassung 8

1. Introduction 10

2. Theoretical background 13

2.1 Justice in schools 13

2.2 The treatment of student misbehavior 16

2.3 Retribution and utilitarianism as goals of punishment 17

2.4 The present research program 19

3. Project 1: “For the greater good? The motivational basis of third-party punishment in children” 24

3.1 Abstract 24

3.2 The present research 27

3.3 Study 1 29

3.3.1 Methods 30

3.3.2 Results and discussion 32

3.4 Study 2 36

3.4.1 Methods 37

3.4.2 Results and discussion 41

3.5 General discussion 44

4. Project 2: “Punishment goals in classroom interventions: An attributional approach” ... 49

4.1 Abstract 49

4.2 The present research 53

4.3 Experiment 1 55

4.3.1 Methods 55
4.3.2 Results and discussion 59
4.4 Experiment 2 64
4.4.1 Methods 65
4.4.2 Results and discussion 66
4.5 Experiment 3 69
4.5.1 Methods 69
4.5.2 Results and discussion 70
4.6 General discussion 73
5. Project 3: "Appropriate but unfair? Comparing teachers' and students' perspectives on the treatment of student misbehavior" 79
5.1 Abstract 79
5.2 The present research 83
5.2.1 Methods 85
5.2.2 Results 90
5.3 Discussion 96
6. General discussion 102
6.1 Theoretical implications 105
6.2 Practical implications 108
6.3 Limitations and future research 111
6.4 Conclusions 113
7. References 115
Acknowledgments 135
Statement of originality 136
Curriculum Vitae 137

LIST OF TABLES

Table 1. Punishment reactions derived from the pretest and used in the main experiments ... 56

Table 2. Means and standard deviations of the endorsement ratings of the three punishment goals ... 91

Table 3. Means and standard deviations of the ratings of the three punishment reactions ... 94

LIST OF FIGURES

Figure 1. Mean rank preference scores of the three punishment goals in each scenario, computed from the order of the questions selected 33

Figure 2. Mean information importance ratings of the three punishment goals in each scenario 35

Figure 3. Coins spent for punishment in the four conditions across participants 42

Figure 4. Results of the direct rating of punishment goals in the three experiments: mean endorsement of punishment goals depending on the attribution of the misbehavior in terms of the stability and controllability of its cause 62

Figure 5. Results of the indirect measurement of punishment goals in the three experiments: mean appropriateness ratings of the three punishment reactions depending on the attribution of the misbehavior in terms of the stability and controllability of its cause 63

Figure 6. Results of the direct rating of punishment goals from the two samples: mean endorsement of the three punishment goals 93

Figure 7. Results of the indirect rating measure of punishment goals from the two samples: teachers' appropriateness ratings of the three punishment reactions, and the mean of students' ratings on how fair, appropriate, and just the teachers' punishment reactions are 96

ABSTRACT

Student misbehavior and its treatment is a major challenge for teachers and a threat to their well-being. Indeed, teachers are obliged to punish student misbehavior on a regular basis. Additionally, teachers' punishment decisions are among the most frequently reported situations when it comes to students' experiences of injustice in school. By implication, it is crucial to understand teachers' treatment of student misbehavior vis-à-vis students' perceptions. One key dimension of punishment behavior reflects its underlying motivation and goals. People generally intend to achieve three goals when punishing misbehavior, namely, retribution (i.e., evening out the harm caused), special prevention (i.e., preventing recidivism of the offender), and general prevention (i.e., preventing imitation of others). Importantly, people's support of these punishment goals is subject to hierarchy and power, implying that teachers' and students' punishment goal preferences differ. In this dissertation, I present three research projects that shed first light on teachers' punishment and its goals along with the students' perception of classroom intervention strategies pursuing these goals. More specifically, I first examined students' (i.e., children's) general support of each of the three punishment goals sketched above. Furthermore, I applied an attributional approach to understand and study the goals teachers intend to achieve when punishing student misbehavior. Finally, I investigated teachers' and students' support of the punishment goals regarding the same student misbehavior to directly compare their views on these goals and reactions pursuing them. In sum, the findings show that students generally prefer retribution and special prevention to general prevention, whereas teachers prefer general prevention and special prevention to retribution. This ultimately translates into a "mismatch" of teachers and students in their preferences for specific punishment goals, and the findings suggest that this may indeed enhance students' perception of injustice. Overall, the results of the present research program may be valuable for the development of classroom intervention strategies that may reduce rather than enhance conflicts in student-teacher-interactions.

ZUSAMMENFASSUNG

Fehlverhalten von SchülerInnen und dessen Klärung ist eine gewaltige Herausforderung für Lehrkräfte und eine große Belastung für deren Wohlbefinden. Lehrkräfte sind regelmäßig gezwungen derartiges Fehlverhalten zu sanktionieren. Gleichzeitig nennen SchülerInnen zumeist die Strafentscheidungen von Lehrkräften, wenn sie über Situationen berichten sollen, bei denen sie in der Schule Ungerechtigkeit erfahren haben. Deshalb ist es wichtig den Umgang von Lehrkräften mit Fehlverhalten in der Schule zu untersuchen und hierbei die Wahrnehmung der SchülerInnen zu berücksichtigen. Ein zentraler Aspekt des Strafverhaltens von Lehrkräften ist dessen zugrundeliegende Motivation und das Ziel der Strafe. Menschen verfolgen generell drei Ziele, wenn sie auf Fehlverhalten reagieren: Retribution (d.h. der Ausgleich des entstandenen Schadens), Spezialprävention (d.h. das Verhindern weiteren Fehlverhaltens durch den Täter oder die Täterin) und Generalprävention (d.h. das Verhindern von Nachahmungstaten durch andere). Hierbei hängen die befürworteten Strafziele vom Status und der Macht der strafenden Person ab, weshalb man erwarten könnte, dass sich die Strafzielpräferenzen von Lehrkräften und SchülerInnen unterscheiden. In dieser Dissertation stelle ich drei Forschungsprojekte vor, in denen ein erster Blick auf die Strafziele von Lehrkräften geworfen wurde. Darüber hinaus wurde die Wahrnehmung von SchülerInnen bezüglich unterschiedlicher Interventionsstrategien im Klassenzimmer und deren Ziele betrachtet. Hierbei habe ich zunächst die generelle Befürwortung der drei oben skizzierten Strafziele durch SchülerInnen (bzw. Kindern) untersucht. Anschließend habe ich auf der Basis von Attributionsmodellen die Strafziele von Lehrkräften analysiert. Zuletzt habe ich die Befürwortung der Strafziele durch Lehrkräfte und SchülerInnen an einem spezifischen Fall von Fehlverhalten im Schulkontext betrachtet und deren Sichtweise auf diese Strafziele, sowie auf Verhalten welches diese Ziele verfolgt, verglichen. Insgesamt zeigen die Ergebnisse, dass SchülerInnen eine Präferenz für Retribution und Spezialprävention gegenüber Generalprävention haben. Lehrkräfte hingegen haben eine Präferenz für Generalprä-

vention und Spezialprävention gegenüber Retribution. Dies führt schlussendlich zu einer Diskrepanz zwischen den Strafzielen die von Lehrkräften und SchülerInnen befürwortet werden. Darüber hinaus fanden sich erste Hinweise, dass diese Diskrepanz zu einer erhöhten Wahrnehmung von Ungerechtigkeit durch die SchülerInnen beiträgt. Alles in allem könnten die Ergebnisse des vorliegenden Forschungsprogrammes einen wertvollen Beitrag zur Entwicklung von Interventionsstrategien im Unterricht leisten, welche Konflikte zwischen Lehrkräften und SchülerInnen nicht verstärken, sondern diese vielmehr reduzieren.

1. INTRODUCTION

In March 2006, a high school in Berlin called "Rütli-Hauptschule" gained wide attention throughout Germany because their teachers published an open letter in the media, disclosing that they considered themselves incapable of continuing to teach their students due to the large extent of student misbehavior. A part of their statement read as follows: "We are helpless, we don't know what to do. We are at the end of a one-way street… There are no values, which guide our school community. There are no role models to guide us, both teachers and pupils. We are isolated and desperate. We are drowning in violence, disrespect and ignorance—the Minister of Education has to secure police protection for our school." (Hainmüller, 2006, p. 115). This inglorious but prominent case shook the German public and was the starting point for vast domestic political debates about the German school system and the problem of student misbehavior in German schools. Unfortunately, this incident remained by no means an exception. More recently, for example, teachers of a German elementary school reported similar problems ("Grundschule in Sachsen-Anhalt", 2018) and recent official statistics show that the frequency of student misbehavior is still increasing (Landeskriminalamt Nordrhein-Westfalen, 2017).

In fact, educational research has also confirmed that a wide range of student misbehaviors occur in class on a daily basis (e.g., Kulinna, Cothran, & Regualos, 2006; Wheldall & Merrett, 1988). These misbehaviors and their treatment have severe consequences for both teachers and students: for teachers, the extent of student misbehavior has found to be strongly linked to their well-being and health, and it has been identified as the most salient stressor related to burnout syndrome in teachers (Aloe, Shisler, Norris, Nickerson, & Rinker, 2014; Brouwers & Tomic, 2000; McCormick & Barnett, 2011). For students, teachers' punishment reactions to classroom misbehavior are a fundamental cause of their experiences of injustice in school (Fan & Chan, 1999; Israelashvili, 1997). Importantly, students' perceptions of injustice in school have a strong impact on their lives, both within the school (e.g., on students' academic

self-concept, motivation, and achievements; Peter, Kloeckner, Dalbert, & Radant, 2012) and beyond (e.g., on students' attitudes towards democracy; Pretsch & Ehrhardt-Madapathi, 2018). Consequently, it is vital to examine and understand teachers' punishment behavior to benefit the development of successful classroom intervention strategies. Furthermore, it is equally important to consider the students' perspectives on teachers' punishment decisions to identify potential characteristics of the punishment that may enhance experiences of injustice.

One aspect of teachers' punishment decisions that is worth examining may be the goals teachers intend to achieve when reacting to student misbehavior. People generally intend to achieve three goals when reacting to misbehavior: retribution (i.e., evening out the harm caused), special prevention (i.e., preventing recidivism of the offender), and general prevention (i.e., preventing imitation of others). Importantly, the support of these goals is found to be subject to power and hierarchy, that is, people differing in power differ in their preferences for specific goals (Mooijman, van Dijk, Ellemers, & van Dijk, 2015). These results may have direct implications for the school context, given that teachers are typically in a more powerful position than students (Goddard, 2000). Consequently, teachers and students may perceive different goals as appropriate in the treatment of student misbehavior. This then may result in undesirable outcomes, such as an increase in norm violations or the experience of injustice in students (Mooijman, van Dijk, van Dijk, & Ellemers, 2017).

Although the study of punishment has a long tradition in the field of social psychology (e.g., Ferster & Skinner, 1957; Skinner, 1938), social psychological research on teachers' treatment of student misbehavior and its goals is still scarce. Therefore, the aim of the present research program was to shed first light on (i.) teachers' punishment and its goals and (ii.) the students' perception of classroom intervention strategies pursuing these goals. In the following, I will give a more detailed overview of the literature on why the study of punishment in school is important and will describe the central concepts of the present research (Chapter 2). Subse-

quently, I will present three research projects prescribed to study potential differences in students' and teachers' perceptions of different punishment goals. More precisely, in project 1, I examined the punishment goal preferences of children around the ages of 10 and 11, given that, so far, all research on the motivational basis of people's punishment was conducted on adults (including the findings the present research is based on) and one may not simply transfer behavioral patterns observed in adults to children (Chapter 3). In project 2, I applied an attributional approach to understand and study teachers' punishment behavior, and in particular, the goals they intend to achieve when reacting to misbehavior (Chapter 4). In project 3, teachers' and students' punishment goal preferences were finally investigated with regard to the same student misbehavior, allowing to compare their views on different punishment reactions and their goals more directly (Chapter 5). To conclude, I will discuss the general results of the three projects, the implications and potential limitations of this research, before drawing overall conclusions (Chapter 6).

2. THEORETICAL BACKGROUND

Justice is considered a fundamental motive in both non-human (Jensen, Call, & Tomasello, 2007) and human primates (Lerner, 1977, 1980). Indeed, throughout humans' evolutionary development from small and insular hunter-gatherer tribes to a large-scale society, justice was and still is a key dimension on both a macro-level (such as between nations; e.g., Wenzel, Lawrence-Wood, Okimoto, & Hornsey, 2017; Wenzel, Okimoto, Hornsey, Lawrence-Wood, & Coughlin, 2017) and a more micro-level (such as in interpersonal contacts or institutions; e.g., Gerlach, Allemand, Agroskin, & Denissen, 2012; Holtz & Harold, 2013).

Correspondingly, the study of institutional justice perceptions and their impact has a long tradition in the field of organizational psychology (e.g., Colquitt, Conlon, Wesson, & Porter, 2001). For example, employees who feel treated unfairly in the workplace report reduced well-being and work satisfaction (Schmitt & Dörfel, 1999), as well as a higher degree of counterproductive work behavior (Ambrose, Seabright, & Schminke, 2002; Furnham & Siegel, 2011). One of the most important organizations in human development and a fundamental socialization setting for individuals in the Western world is the school environment (Pretsch & Ehrhardt-Madapathi, 2018; Resh & Sabbagh, 2014). Consequently, research examining the role of justice in school (i.e., its influences and consequences) is of great importance.

2.1 JUSTICE IN SCHOOLS

Children spend a great amount of their time at school and, in everyday classroom interactions, the issue of justice is ubiquitous: for example, students may perceive to be treated unfairly by classmates, for instance, as being excluded during breaks. Likewise, students may also perceive to be treated unfairly by teachers, for instance, concerning the grading of tests or the teacher's attention during the lesson. Correspondingly, scholars in the domain of educational research have identified the topic of justice as an important aspect in school (Sabbagh & Resh, 2016). For example, this research has revealed that the perception of justice (vs. injustice) is

associated with a better class climate (Peter & Dalbert, 2010), less bullying behavior (Donat, Knigge, & Dalbert, 2018; Donat, Umlauft, Dalbert, & Kamble, 2012), and less student aggression in general (Chory-Assad, 2002). Likewise, heightened experiences of injustice in school reduce students' well-being and enhance their level of distress (Dalbert & Stoeber, 2005; Kamble & Dalbert, 2012), increase school absenteeism (Donat, Gallschütz, & Dalbert, 2018), and, ultimately, decreases students' academic self-concept, motivation, and achievements (Peter et al., 2012).

Given the focus of teachers and students on school achievements and the increasing performance pressure in most Western societies, one may expect that the assessment of student performances (i.e., the grading of tests) is of primary importance for students' perceptions of injustice (Dalbert, Schneidewind, & Saalbach, 2007). However, students only scarcely list incidents of grading or the evaluation of their performances when asked to describe situations in which they felt treated unfairly in school (Israelashvili, 1997). By contrast, it is the treatment of student misbehavior that is among the most frequently reported situations when it comes to students' experiences of injustice, with situations of false allegations not even included (Fan & Chan, 1999). In other words, in situations in which the misbehavior itself appears to be clear, students frequently perceive the punishment reaction of teachers as unjust.

However, it is not only the students for whom classroom misbehavior and its treatment by teachers is unsatisfying and problematic. The occurrence of student misbehavior is also positively associated with many health-related problems in teachers, most prominently issues regarding satisfaction, well-being, stress and, ultimately, burnout (Aloe et al., 2014; Boyle, Borg, Falzon, & Baglioni, 1995; McCormick & Barnett, 2011). Additionally, many teachers indicate a lack of self-confidence in dealing with such incidents and express a desire for support and guidance (e.g., through training or supervision; Chan, 1998; Ingersoll, 2001; Melnick & Meister, 2008).

Although the issue of student misbehavior has received much attention within the last decades (Aloe et al., 2014; Infantino & Little, 2005), there is no single definition of student misbehavior in the literature. Most works conceptualize it as a behavior that disrupts the teaching and learning processes, impairs the operation of the class, or violates general societal rules or norms (e.g., Finn, Fish, & Scott, 2008). Examples of student misbehavior include disrupting instruction, noncompliance or failure to follow directions, verbal abuses or disrespect toward teachers, bullying, and physical or verbal abuse towards class mates (Brouwers & Tomic, 2000; Bru, Stephens, & Torsheim, 2002; Finn et al., 2008; Kulinna et al., 2006).

The most frequently reported student misbehaviors are of rather mild format, such as talking out of turn, hindering others, and idleness or slowness (Infantino & Little, 2005; Reynolds, Stephenson, & Beaman, 2011; Wheldall & Merrett, 1988). However, even those minor forms of student misbehavior represent a fundamental challenge for most teachers (Little, 2005; Ratcliff, Jones, Costner, Savage-Davis, & Hunt, 2010; Wheldall & Merrett, 1988). Additionally, although such student misbehavior may—in the best case—be prevented by proactive classroom management approaches (e.g., through supporting positive behaviors; Sugai & Horner, 2006), teachers are also forced to react to all other forms of misconduct at school that may not be prevented by a well-prepared lesson (e.g., if it occurs during break). Therefore, in some situations, the classroom may bear a resemblance to a courtroom, with the teacher representing the judge who is in charge of finding an appropriate punishment (Weiner, 2003).

The occurrence of student misbehavior and teachers' classroom interventions can best be described as a dynamic process (Aloe et al., 2014; Ratcliff et al., 2010). This process depends on the teacher's appraisal of the behavior, their perceived self-efficacy, and the strategies applied to react to the behavior (Chang, 2009; Lewis, 1999; Ringer, Doerr, Hollenshead, & Wills, 1993). Teachers' management of student misbehavior, then, not only influences the misbehaving student herself (Bru et al., 2002), but also the classroom climate and may, in the worst case, even result in a bad relationship between students and teacher (Avtgis & Rancer, 2008; Ratcliff

et al., 2010). In turn, a tense student-teacher relationship and the students' perceptions of low support from the teacher may increase conflict and, eventually, the occurrence of further classroom misbehavior (Boyle et al., 1995; Bru, Murgberg, & Stephens, 2001; Bru et al., 2002; Ertesvåg & Vaaland, 2007). Therefore, the "wrong" reaction of teachers to student misbehavior may even cause higher levels of negative attitudes among students towards learning, and it may thereby enhance further misconducts (Lewis, Romi, Katz, & Qui, 2008). Thus, it is crucial to understand and study teachers' punishment behavior to identify characteristics that do not help but harm the proficient operation of the class.

2.2 THE TREATMENT OF STUDENT MISBEHAVIOR

Teachers have a considerable amount of freedom to react to undesirable student behaviors. Classic intervention strategies in the school are public or private praise and reprimand, sending the offender to the principal's office or enforcing a period of time out, extra recess or losing recess for the class or for the offender, and point reward systems (Andreou & Rapti, 2010; Elliott, 1986, 1988; Kulinna, 2008; Turco & Elliott, 1986).

A majority of research on these classroom management techniques has focused on their effectiveness in operating the class and decreasing the extent of student misbehavior (e.g., Lewis, Romi, Qui, & Katz, 2005; Little, Hudson, & Wilks, 2002; Von Brock & Elliott, 1987). In turn, only few research has also considered the students' perspective on different classroom intervention strategies to potentially identify characteristics that may be associated with their approval or rejection of such teacher behavior (Elliott, 1986; Lewis, 2001; Lewis et al., 2008; Lovegrove, Lewis, Fall, & Lovegrove, 1985; Turco & Elliott, 1986).

However, most research examining both teachers' and students' perspectives on the punishment of student misbehavior only compared their preferences regarding a set of specific classroom discipline practices (e.g., Lewis, 2001; Lovegrove et al., 1985). By contrast, the psychological mechanisms underlying students' perception and evaluation of punishment in school

are largely unknown. For example, whereas some classroom intervention strategies focus on treating the misbehaving student in particular, others implement forms of collective punishment (Elliott, 1988). Unsurprisingly, students show a preference for the specific treatment of the actually misbehaving student over collective punishment procedures (Elliott, Witt, Galvin, & Moe, 1986). One other characteristic of teachers' punishment reactions that may influence the students' perception of teachers' punishment decisions is the goals teachers intend to achieve when reacting to student misbehavior (Carlsmith, Darley, & Robinson, 2002).

2.3 RETRIBUTION AND UTILITARIANISM AS GOALS OF PUNISHMENT

There is considerable social psychological research on laypeople's punishment goals (Carlsmith et al., 2002; Cushman, 2015; Goodwin & Gromet, 2014). This research is inextricably associated with the philosophical works by Immanuel Kant and Jeremy Bentham. According to Kant (1952), punishment should primarily be motivated by the strive to rebalance the (moral) wrong that has been committed by the offense. This punishment philosophy is referred to as *retribution* (Carlsmith & Darley, 2008; Carlsmith et al., 2002). A retributive punishment should aim at paying back harm doers for their misconduct and, thus, to restore justice (Gerber & Jackson, 2013). This is typically achieved by finding a proportionate punishment to retaliate the (intended) harm caused by the wrong committed (Goodwin & Gromet, 2014). Therefore, the punishment is ought to "fit the crime" (Carlsmith, 2006) and, hence, retribution is also referred to as "the just deserts" theory (Carlsmith et al., 2002).

A normatively different approach of punishment is strongly related to the ideas of Jeremy Bentham (1962), who considered the act of punishment from a more *utilitarian* view on people's behavior. Correspondingly, this philosophy perceived the intrinsically damaging act of punishment merely as justifiable if it leads to positive consequences—in particular, by preventing future misbehavior (e.g., McCullough, Kurzban, & Tabak, 2013). Importantly, utilitarian punishment can further be differentiated into *special prevention* and *general prevention*

(Goodwin & Benforado, 2015; Keller, Oswald, Stucki, & Gollwitzer, 2010; Rucker, Polifroni, Tetlock, & Scott, 2004). A special preventive punishment is primarily concerned with the offender herself and intends to prevent future recidivism (Keller et al., 2010). A general preventive punishment is, in turn, primarily concerned with other members of the community that might have been informed of the offense and, therefore, may imitate the misbehavior if it remains unpunished (Goodwin & Benforado, 2015).

A great body of research studying the motivational basis of people's punishment has dealt with the question whether punishment behavior is primarily driven by either retributive or utilitarian purposes—with a seemingly clear answer: punishment, it appears, is predominantly driven by retribution (i.e., to even out the wrong that has been done) rather than special or general prevention (i.e., to prevent future crimes by the offender or others; Carlsmith, 2006; Keller et al., 2010). This has led to the consensus that "people are intuitive retributivists" (Carlsmith & Darley, 2008, p. 211).

However, more recently, there is increasing skepticism about the ubiquity of the "intuitive retributivists" position (Goodwin & Gromet, 2014), with research highlighting the role of preventive goals on individuals' actual punishment behavior (e.g., Crockett, Özdemir, & Fehr, 2014). This research has also started to identify potential moderators of people's punishment goals, such as the punisher's personality (Giacomantonio & Pierro, 2014; Giacomantonio, Pierro, Baldner, & Kruglanski, 2017) or the attribution of the misbehavior in terms of the controllability and stability of its cause (Graham, Weiner, & Zucker, 1997; Weiner, Graham, & Reyna, 1997).

Important for the present research, recent studies have revealed that hierarchy and power also affect people's support of different punishment goals. More precisely, it has been found that leaders (i.e., people in a powerful position) prefer utilitarian punishment, whereas subordinates prefer retributive punishment (Mooijman et al., 2015). These results may have direct implications for the school context. That is, in school, teachers are typically in a more powerful

position and may, thus, be considered "leaders" in the school context (Goddard, 2000). Whereas, students are typically in a less powerful position in school and, thus, may be considered "subordinates" in the school context. Consequently, teachers and students may perceive different goals as appropriate in the treatment of student misbehavior. Stated differently, the differences in status and power between teachers and students may lead to an inherent "mismatch" in their punishment goal preferences, which in turn, may increase further misconducts and enhance students' experiences of injustice (Mooijman et al., 2017).

2.4 THE PRESENT RESEARCH PROGRAM

In the present research, I provide a first analysis of the goals teachers intend to achieve when reacting to student misbehavior. Furthermore, I examine the students' perception of classroom intervention strategies pursuing different goals. In particular, I suggest that teachers and students have different punishment goal preferences leading to a "goal mismatch" that may, ultimately, cause students' perception of injustice in the treatment of misbehavior.

So far, all research examining individuals' punishment goals has been conducted with adult participants (such as the aforementioned research on power and punishment goals the present research is based on). Importantly, however, one cannot simply transfer behavioral patterns observed in adults to children or adolescents, given that moral judgment and justice-related decision making is known to be subject to an ontogenetic development (Carpendale, 2000; Killen & Smetana, 2014, 2015; Kohlberg, 1981). Although there is considerable evidence showing that children engage in punishment even if they cannot expect any future benefits from doing so (Jordan, McAuliffe, & Warneken, 2014; McAuliffe, Jordan, & Warneken, 2015), studies examining the motivational basis of children's punishment behavior are entirely missing. Therefore, the first project of this dissertation is dedicated to study children's punishment goals. More precisely, I examine whether children—in the position of "subordinates" in the theoretical

framework underlying the present research (Mooijman et al., 2015)—show the expected preferences in their punishment goals, that is, preferring retributive over utilitarian punishment.

Moreover, it is not only important to study the children's perspective on punishment goals but also to investigate the actual goals of teachers reacting to student misbehavior. Importantly, in contrast to participants in most social psychological studies on punishment behavior (again, such as the research on power and punishment goals the present research is based on), teachers in charge of a punishment decision at school typically have considerable background information about the specific individuals involved in the misconduct (i.e., the offender or a potential victim). Such background information (e.g., on the offender's character) may have a fundamental impact on the perception of the misbehavior (Uhlmann, Pizarro, & Diermeier, 2015) and, ultimately, the teacher's punishment goals (Gromet & Darley, 2009a). In particular, teachers' punishment goal preferences may be subject to the perceived stability and controllability of the cause of the misbehavior (Graham et al., 1997; Weiner et al., 1997). Consequently, to receive a comprehensive picture of teachers' punishment behavior and its goals, it is crucial to examine teachers' punishment in varying situations of student misbehavior (i.e., varying in the attribution of the misbehavior). This is the aim of the second project of this dissertation. More precisely, in this project, I examine whether teachers—in the position of "leaders" in the theoretical framework underlying the present research (Mooijman et al., 2015)—show the expected preferences in their punishment goals, that is, preferring utilitarian over retributive punishment. Furthermore, I examine whether these preferences are subject to the attribution of the specific student misbehavior.

Lastly, a key objective of the present research program is to directly compare students' and teachers' perceptions and their support of different punishment goals. Consequently, it is necessary to provide both students and teachers with the same instance of student misbehavior and assess their support of different punishment goals in this particular situation. Importantly, it is vital to examine the goals teachers and students generally support in the given situation, as

well as their evaluation of different punishment reactions a teacher could show—with the reactions differing in the goals they achieve. This is the topic of the third project of this dissertation. More precisely, I examine whether students and teachers show the expected "mismatch" in their punishment goal preferences (i.e., with students preferring retributive over utilitarian punishment; and teachers preferring utilitarian over retributive punishment).

In the present research program, several aspects are worth highlighting. First, I consistently apply various measures to assess participants' punishment goals and, thus, do not exclusively rely on self-reports on the goals participants explicitly endorse. This is important given that past research has revealed a rather weak correlation between people's explicit endorsement of punishment goals (i.e., as provided in self-reports) and the goals their actual punishment behavior is aiming at. In particular, people tend to support *all* goals of punishment and indicate a considerable *endorsement* of utilitarian punishment, whereas their actual punishment *behavior* is mostly retributive (Carlsmith, 2008; Crockett et al., 2014). Consequently, it is important to consider both self-report measures of goal endorsement, as well as more indirect ways to assess people's punishment goals that do not require individuals to introspectively predict the exact goals they might intend to achieve. Thus, in all projects, I apply a range of other methods to assess participants' punishment goals in addition to explicit goal endorsement measures (e.g., an information-selection task and an economic game in project 1; indirect reaction measures in project 2 and 3).

Second, as outlined above, the concept of (student) misbehavior comprises a large amount of different types of behaviors. These misbehaviors can vary on several dimensions that affect individuals' punishment behavior, such as the severity of the offense (Gromet & Darley, 2009b; Rucker et al., 2004) or the role of the punisher in the offense situation. More specifically, in a situation of misbehavior, individuals can take various perspectives: the role of the offender, the role of the victim, or the role of an observer (Schmitt, Gollwitzer, Maes, & Arbach, 2005). Research on people's punishment behavior and its goals mostly examines two roles a punisher

can take: the punisher can take the position of the victim (also referred to as second-party punishment) or of a (most often unaffected) bystander (also referred to as third-party punishment). Indeed, the position in the offense situation is essential for individuals' punishment decisions (FeldmanHall, Sokol-Hessner, Van Bavel, & Phelps, 2014). Clearly, both perspectives are of great interest in general and in the specific context of classroom misbehavior. However, the most frequently reported student misbehaviors are not directly aimed against the teacher (i.e., a situation requiring second-party punishment; Infantino & Little, 2005; Reynolds et al., 2011). Therefore, the present research entirely focuses on third-party punishment situations (i.e., participants observed a misbehavior of an individual towards another person) and does not consider any situations of second-party punishment.

Lastly, most research on people's punishment goals focused on the dichotomy between retribution and utilitarianism (Carlsmith et al., 2002; Crockett et al., 2014; Mooijman et al., 2015)—with the latter representing the broad goal of preventing future misconducts, sometimes specified as preventing recidivism (i.e., special prevention; Crockett et al., 2014), and sometimes specified as preventing imitation (i.e., general prevention; Mooijman et al., 2017). However, especially in the context of the punishment of student misbehavior, it is arguably important to differentiate between special prevention and general prevention, given that student misconducts often occur publicly: they are observed by—and may even directly affect—other students. Therefore, teachers should be generally interested in preventing both further misconduct by the offending student (i.e., special prevention) and imitation by observers (i.e., general prevention). However, in some situations, teachers may be interested in merely pursuing one of these utilitarian goals; for example in a situation in which the misbehavior and punishment may be extremely shameful for the student and a general preventive (i.e., public) punishment could cause severe future consequences, such as the social exclusion from classmates. Hence, across all

three projects, I divided the utilitarian punishment goal into special prevention and general prevention as two separate forms and, thus, examined participants' support of retribution, special prevention, and general prevention.

3. PROJECT 1: "FOR THE GREATER GOOD? THE MOTIVATIONAL BASIS OF THIRD-PARTY PUNISHMENT IN CHILDREN"

3.1 ABSTRACT

People willingly accept personal costs to sanction norm violations even if they are not personally affected by the wrongdoing and although their sanctioning may not yield any future benefits—a behavior known as third-party punishment. Strikingly, this behavior has also been observed in young children. However, little is known about the motives underlying third-party punishment in children. In two studies, we applied different methodological approaches measuring the punishment goals pursued by children around the ages of 10 and 11. Specifically, we were interested in the extent to which their punishment was driven by retribution (i.e., evening out the harm caused), special prevention (i.e., preventing recidivism of the offender), and general prevention (i.e., preventing imitation of others). In Study 1 (N = 69), children were both asked to select from information provided about two misbehaviors before indicating a punishment, and to rate the importance of each piece of information for their punishment. Importantly, these information pieces are known to be associated with different punishment goals. In Study 2 (N = 238), we adapted a third-party punishment game measuring isolated punishment motives by controlling whether the offender (special prevention), a bystander (general prevention), or nobody (retribution) learns about the punishment. We consistently found that children's punishment is predominantly motivated by both retributive and preventive purposes. Specifically, special prevention appeared to be as important as retribution whereas general prevention was only of importance in the economic game. These results contribute to an increasing body of research expressing criticism on the idea of people as "intuitive retributivists" and provide practical implications for the treatment of child misbehavior.

For the greater good? The motivational basis of third-party punishment in children

People willingly sacrifice own resources (e.g., time or money) to sanction norm violations across a wide range of societies (Henrich et al., 2006). Strikingly, many individuals accept personal costs to punish offenders even if they are not affected by the wrongdoing, and even if their sanctioning does not yield any future benefits (Fehr & Gächter, 2002). In social psychological research, this "altruistic act" of moral courage is referred to as *third-party punishment* (Boyd, Gintis, Bowles, & Richerson, 2003; Fehr & Fischbacher, 2004).

For over a decade, researchers have sought to understand the underlying motives driving this behavior (e.g., Balafoutas, Nikiforakis, & Rockenbach, 2014; Bowles & Gintis, 2004; Boyd, Gintis, & Bowles, 2010; Carlsmith, 2006; Carlsmith et al., 2002; Gardner & West, 2004; Gintis, Bowles, Boyd, & Fehr, 2003; Keller et al., 2010). This research has identified two main motives for third-party punishment: *utilitarianism* and *retribution*. On the one hand, according to a utilitarian approach, punishment is an instrument to facilitate compliance with social norms, leading to more cooperative behavior and decreasing norm violations (Fehr & Fischbacher, 2004; Henrich et al., 2006). Thus, according to this approach, punishment behavior is primarily forward-oriented and aims to prevent further misconduct, either by the perpetrator herself (termed *special prevention*) or by unrelated observers (termed *general prevention*; Keller et al., 2010; Rucker et al., 2004). On the other hand, according to a retributive approach, punishment is intended to rebalance the moral wrong caused by the offense and seeks to pay back transgressors for their misbehavior (Carlsmith, 2006; Carlsmith et al., 2002; Goodwin & Gromet, 2014). Thus, according to this approach, punishment behavior is primarily backward-oriented and concerned about the harm caused but not about future developments (Carlsmith et al., 2002).

Understanding laypeople's opinions on the purposes of punishment is crucial, as the willingness to comply with rules highly depends on the perceived legitimation of such principles (Nadler, 2005). One aspect of this legitimation is indeed the goal an authority intends to achieve when punishing a wrongdoing. For example, it has been found that the pursuit of utilitarian punishment goals by authorities actually leads to a decrease in compliance to rules and social norms by subordinates (Mooijman et al., 2017), which may be due to differences between the expected (and presumably legitimate) and the actually pursued punishment goals.

A large body of research has analyzed people's punishment decisions to test whether there is a primary objective underlying such behavior—with the result that, overall, punishment is predominantly driven by retribution (i.e., to even out the harm caused) rather than special or general preventive purposes (i.e., to prevent further misconduct by the offender or others; Carlsmith, 2006, 2008; Carlsmith et al., 2002). This led to an apparent consensus that "people are intuitive retributivists" (Carlsmith & Darley, 2008, p. 211), although more recent research suggests that the ubiquity of the "intuitive retributivists" position may, at least partially, be caused by methodological reasons (e.g., Goodwin & Benforado, 2015).

Whereas much research investigated the motives underlying altruistic third-party punishment of adults, far less is known about children's punishment goals. This lack of evidence is striking for at least two reasons. First, the perceived legitimacy of punishment is particularly crucial for children (Tisak, 1986) and indeed an almost day-to-day experience in contexts such as school (Kulinna et al., 2006). Second, there is considerable evidence showing that children themselves engage in third-party punishment. More specifically, it has been shown that young children understand the function and importance of social norms and rules already by age 3 (Schmidt, Rakoczy, & Tomasello, 2012; Smith, Blake, & Harris, 2013). They express a general preference of equal shares (Hamann, Warneken, Greenberg, & Tomasello, 2011; Lobue, Nishida, Chiong, Deloache, & Haidt, 2011; Schmidt & Sommerville, 2011) and, at least by age

4, deprecate unfair treatment of individuals, independent of whether they themselves are disadvantaged (Blake & McAuliffe, 2011; Fehr, Bernhard, & Rockenbach, 2008; Shaw & Olson, 2012). Furthermore, already young children engage in punishment when they themselves have been harmed (Gummerum & Chu, 2014; Robbins & Rochat, 2011) and are willing to intervene and correct behavior that harms others (Hamlin, Wynn, Bloom, & Mahajan, 2011; Kenward & Östh, 2012; Riedl, Jensen, Call, & Tomasello, 2015). In fact, by age 6, children start to show altruistic third-party punishment, that is, a costly punishment of norm violations, even if they cannot expect any future benefits from doing so (Jordan et al., 2014; McAuliffe et al., 2015).

Despite these findings, the goals children intend to achieve with their punishment are largely unknown. However, understanding children's motivational basis of third-party punishment is crucial, given the argument stated above: it is arguably children rather than adults who frequently experience misbehavior and punishment (as either a perpetrator, a victim, or an observer), for instance in school (e.g., Kulinna et al., 2006; Wheldall & Merrett, 1988). Such incidents are a main stressor for children (Fan & Chan, 1999; Israelashvili, 1997) that may have a fundamental impact on their lives (Wald & Losen, 2003). Furthermore, differences between the expected (i.e., derived from one's own punishment goals) and actually pursued punishment goals of authorities (e.g., teachers or parents) may lead to a decreasing compliance to rules and norms by the children (Mooijman et al., 2017). Therefore, it is of fundamental importance to understand children's perspective on the treatment of misbehavior (e.g., for schools or families).

3.2 THE PRESENT RESEARCH

We conducted two studies investigating punishment goals of children around the ages of 10 and 11 using three different methodological approaches. None of them require children to introspectively predict the goals they might intend to achieve when punishing an offender,

given that such approaches have already been shown to be biased even for adults, as they generally overestimate the influence of utilitarian goals on their own punishment decisions (Carlsmith, 2008; Crockett et al., 2014). We therefore applied an *information-selection task* (Keller et al., 2010) and an *information importance rating task* (i.e., a task based on the general idea of the information-selection task but without forcing participants to choose between options) in Study 1. In Study 2, we benefit from more recent advances in developmental psychology, borrowing approaches from behavioral decision sciences in the form of economic games (Gummerum, Hanoch, & Keller, 2008). Importantly, we study both retribution and utilitarianism as potential punishment goals of children, but further differentiate the latter into special prevention and general prevention, given that both are distinct as well as essential goals of punishment (Goodwin & Benforado, 2015; Keller et al., 2010; Rucker et al., 2004; Tetlock, 2002).

Given that our research is the first to address the motivational basis of children's third-party punishment, several different outcomes were possible—each with its own theoretical implications. First, children's punishment may primarily be driven by retribution rather than by utilitarian purposes. This would be consistent with research showing a predominance of retribution in driving punishment behavior in adults and would further strengthen the idea of people as "intuitive retributivists" (Carlsmith & Darley, 2008, p. 211). Second, both retribution and the prevention of future misbehavior may be comparably important for children's punishment behavior. This would highlight the importance of punishment decisions frequently perceived in children's daily lives (e.g., in schools, given that teachers' behavior in school is found to be extensively affected by utilitarian purposes; Reyna & Weiner, 2001). Third, children's punishment behavior may be primarily driven by utilitarian punishment goals rather than by retribution, largely questioning the "intuitive retributivists" approach and adding to an increasing body of research expressing criticism on this position (e.g., Goodwin & Benforado, 2015).

For both studies, we report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures (Simmons, Nelson, & Simonsohn, 2012). All materials, the data, and analyses of the two studies are available online at the Open Science Framework (OSF; accessible during double-blind peer-review via the following link: https://osf.io/dwncr/?view_only=b3583932c3e24c76aae729a92c097bb8).

3.3 STUDY 1

In this study, we first applied an *information-selection task* that has already been used in previous research to investigate adults' punishment goals (Keller et al., 2010). In this scenario task, participants are asked to punish an offender for a specified misbehavior. However, before they determine their punishment, participants have the opportunity to request pieces of information about the situation of the misbehavior (e.g., about the magnitude of harm done, the public awareness of the crime, or the offender's motivation for the misbehavior). Each piece of information is a priori known to be largely associated with either of the three punishment goals (Carlsmith, 2006; Keller et al., 2010). Thus, selecting information associated with retribution indicates a higher interest in retributive punishment; selecting information associated with special prevention indicates a higher interest in special preventive punishment; and selecting information associated with general prevention indicates a higher interest in general preventive punishment. This selection task has the advantage that it prevents ceiling effects—a problem that was anticipated given past work on laypeople's punishment goals (Reyna & Weiner, 2001). Additionally, we created an *information importance rating task* by modifying the procedure of the information-selection task while retaining the general idea of "indirectly" measuring participants' punishment goals through their interest in information about the misbehavior. However, in the information importance rating task, participants are not asked to select information about the crime to determine a punishment but to rate the importance of the information provided for their punishment decision. This measurement of punishment goals bears the advantage that it

does not involve a forced choice task (that inherently results in rank-level data). Instead, as a rating scale measure, it allows for use of generally more powerful statistical tests while retaining the indirect nature of measuring punishment goals.

3.3.1 Methods

Measures and procedure

Data collection took place in a German public school. On arrival in the classroom, children provided the experimenters with the consent form signed by their parents or legal guardians. Then, children were seated individually in front of computers. Two experimenters welcomed the class and gave them verbal instructions on the subsequent tasks. The first task was the information-selection task, for which children were told that they will have to determine a punishment for several forms of misbehavior. These misbehaviors were presented in short comic strips. Before learning about the misbehaviors, they were introduced to a superhero called "Superflash," with the superpower of looking into the future whom they would be asking for information. Using an exemplary misbehavior, children received comprehensive instructions on the procedure of the task. After the example, children were asked to work on the subsequent task individually. They were then given two comic strips with different misbehaviors that were presented in randomized order. One comic strip described a misbehavior from a child (pushing another child); the other strip described a misbehavior from an adult (tripping up a waitress). Following each misbehavior, children were provided with a chance to ask Superflash three out of six questions in order to determine their punishment decision. These questions were associated with either retribution, special prevention, or general prevention as punishment goals (with two questions for each goal; Carlsmith, 2006; Keller et al., 2010). Children were instructed to select the questions in the order of their importance (i.e., selecting the most important item first, then the second important item, and so on). After selecting three questions, they received the answers to these questions all at once and were asked to indicate their punishment decision (note that this measure is not of interest for the present research question). After the

information-selection task, children provided demographic information and then reached a stop sign asking them to wait until everyone else had finished part one of the study.

In part two, the experimenters again gave verbal instructions on the second task using the same example as in the beginning. In this part of the study, children were asked to work on the information importance rating task. Specifically, they were presented with the same cases of misbehavior as before, but were now asked to *rate* the importance of the six questions provided to determine a punishment (rather than *selecting* three of these questions). Each question was rated on a 7-point scale ranging from 1 = "*not important*" to 7 = "*very important.*" Again, children worked on this task individually, until reaching a stop sign.

In the final part of the study, children were verbally instructed to give answers on several other variables (not pertinent to the present research question), before they were thanked and verbally debriefed. The full material (including the comic strips used) of this study can be found at the OSF.

Sample

To determine the required sample size, we conducted an a priori power analysis using G*Power (Faul, Erdfelder, Buchner, & Lang, 2009; Faul, Erdfelder, Lang, & Buchner, 2007). Specifically, we aimed to detect a medium-sized effect of $f = .20$ in a repeated measures ANOVA (number of groups = 1, number of measurements = 3 for the three punishment goals to be rated) with a power of $1-\beta = .80$, given a conventional $\alpha = .05$, and nonsphericity correction $\varepsilon = 1$. This resulted in an aspired sample size of $N = 58$ participants. However, because we planned to use Friedman's ANOVA to analyze the data from the information-selection task (Carlsmith, 2006; Keller et al., 2010), we followed the recommendation of Lehmann (2006) and added 15% to the calculated sample size, resulting in a required sample size of $N = 67$ participants. Furthermore, a pragmatic approach was utilized before starting data collection, given that we collaborated with schools for this project. This prevented us from ceasing data

collection exactly upon reaching the required sample size. Therefore, we planned to collect data in one school and to recruit further schools for participation if we did not reach the threshold of approximately 70 participants.

We collected data in the fifth grade of a public German school. After four sessions, N = 69 children participated in the study. Around 39% of participants (i.e., n = 27) were female, most children's mother language was German (91%), and ages ranged between 9 and 13 (M = 11.07, SD = 0.65).

3.3.2 Results and discussion

Preliminary data treatment

For the information-selection task, we computed a rank preference score to assess the relative importance of each punishment goal, closely following the procedure suggested in past research (Carlsmith, 2006; Keller et al., 2010). Thus, each goal received a total preference score depending on the importance of the questions associated with this goal (i.e., depending on the order in which the questions were selected). A question selected first received three points for the respective punishment goal, a question selected second received two points, and a question selected third received one point. For example, if a participant selected both special preventive questions first and a retributive question third, special prevention received a total score of 5 (= 3 + 2), retribution received a total score of 1, and general prevention received a total score of 0. For the information importance rating, we computed the mean importance rating of the two items associated with each punishment goal, respectively. Thus, we received one importance rating for each goal.

Results of the information-selection task

The mean rank preference scores of the punishment goals computed from the information-selection task are shown in Figure 1. For both scenarios, items associated with retribution (M_{Adult} = 2.81, SD_{Adult} = 1.41, and M_{Child} = 2.84, SD_{Child} = 1.36, respectively) received the

highest scores, followed by the special preventive (M_{Adult} = 2.20, SD_{Adult} = 1.33, and M_{Child} = 1.93, SD_{Child} = 1.40, respectively) and general preventive (M_{Adult} = 0.99, SD_{Adult} = 1.23, and M_{Child} = 1.23, SD_{Child} = 1.16, respectively) items. Thus, the general preventive items were perceived as least important.

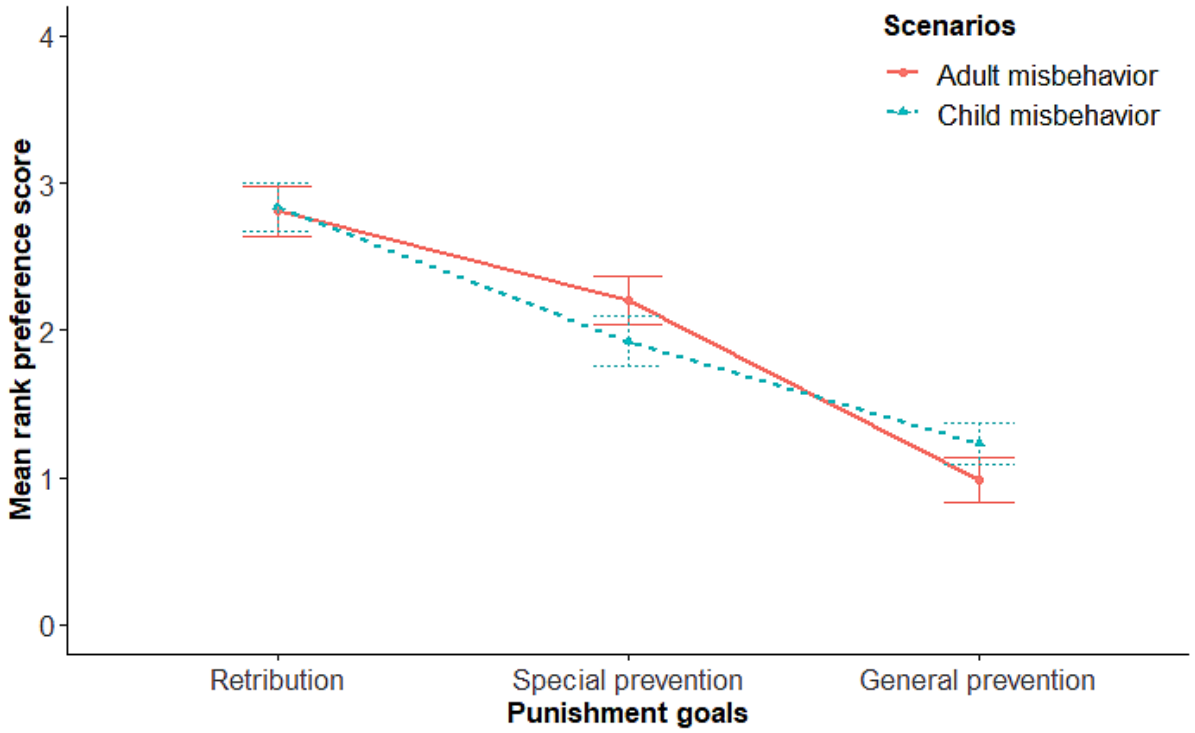

Figure 1. Mean rank preference scores of the three punishment goals in each scenario, computed from the order of the questions selected. Error bars represent one standard error of the mean.

To statistically test this pattern, we submitted the mean rank preference scores of each scenario to a separate Friedman's ANOVA[1], followed by Wilcoxon's signed rank tests as post-hoc pairwise comparisons. In the case of adult misbehavior, the overall ANOVA indicated significant differences between punishment goals, $\chi^2(2) = 42.11$, $p < .001$. Post-hoc tests indicated

[1] Conducting Friedman's ANOVA, it is not possible to include the data of both scenarios in one analysis (considering the scenario as a further factor). Therefore, we conducted two separate analyses, although one could have expected similar results given the descriptive patterns.

significantly higher scores for retribution than special prevention, $z = 2.04$, $p = .041$, $r = .17$, and general prevention, $z = 5.22$, $p < .001$, $r = .44$. Furthermore, special prevention received significantly higher scores than general prevention, $z = 4.01$, $p < .001$, $r = .34$.

Likewise, for the child misbehavior, the overall ANOVA indicated significant differences between punishment goals, $\chi^2(2) = 28.17$, $p < .001$. Similar to the adult case, post-hoc tests indicated significantly higher scores for retribution than special prevention, $z = 2.92$, $p = .004$, $r = .25$, and general prevention, $z = 5.17$, $p < .001$, $r = .44$. Again, special prevention received significantly higher scores than general prevention, $z = 2.49$, $p = .013$, $r = .21$.

Results of the information importance rating task

The mean information importance rating of each punishment goal is displayed in Figure 2. In total, children rated all information as rather important for their punishment decision, leading to overall ceiling effects that have also been observed in past work on laypeople's punishment goals (Reyna & Weiner, 2001). Similar to the results achieved from the information-selection task, for both scenarios, items associated with retribution ($M_{Adult} = 5.80$, $SD_{Adult} = 1.41$, and $M_{Child} = 5.68$, $SD_{Child} = 1.48$, respectively) were rated as most important. However, on this measure, items associated with special prevention ($M_{Adult} = 5.73$, $SD_{Adult} = 1.28$, and $M_{Child} = 5.61$, $SD_{Child} = 1.35$, respectively) were rated as almost equally important. In line with the information-selection task, items associated with general prevention ($M_{Adult} = 4.72$, $SD_{Adult} = 1.88$, and $M_{Child} = 4.80$, $SD_{Child} = 1.75$, respectively) were rated as least important.

We submitted the mean information importance ratings of both scenarios to a repeated measures ANOVA with the punishment goals (retribution, special prevention, general prevention) and the scenarios (adult misbehavior, child misbehavior) as within-subjects factors. This analysis revealed significant differences between punishment goals, $F(2, 136) = 33.43$, $p < .001$, $\hat{\eta}^2_G = .13$. Importantly, there was neither a main effect of the scenario, $F(1, 68) = 0.53$, $p = .469$, $\hat{\eta}^2_G < .001$, nor a significant interaction of the scenario and the punishment goals, $F(2, 136) =$

0.90, p = .408, $\hat{\eta}^2_G$ = .002. Consequently, we aggregated (that is, averaged) ratings of each punishment goal across the two scenarios to conduct pairwise post-hoc *t*-tests between goals.

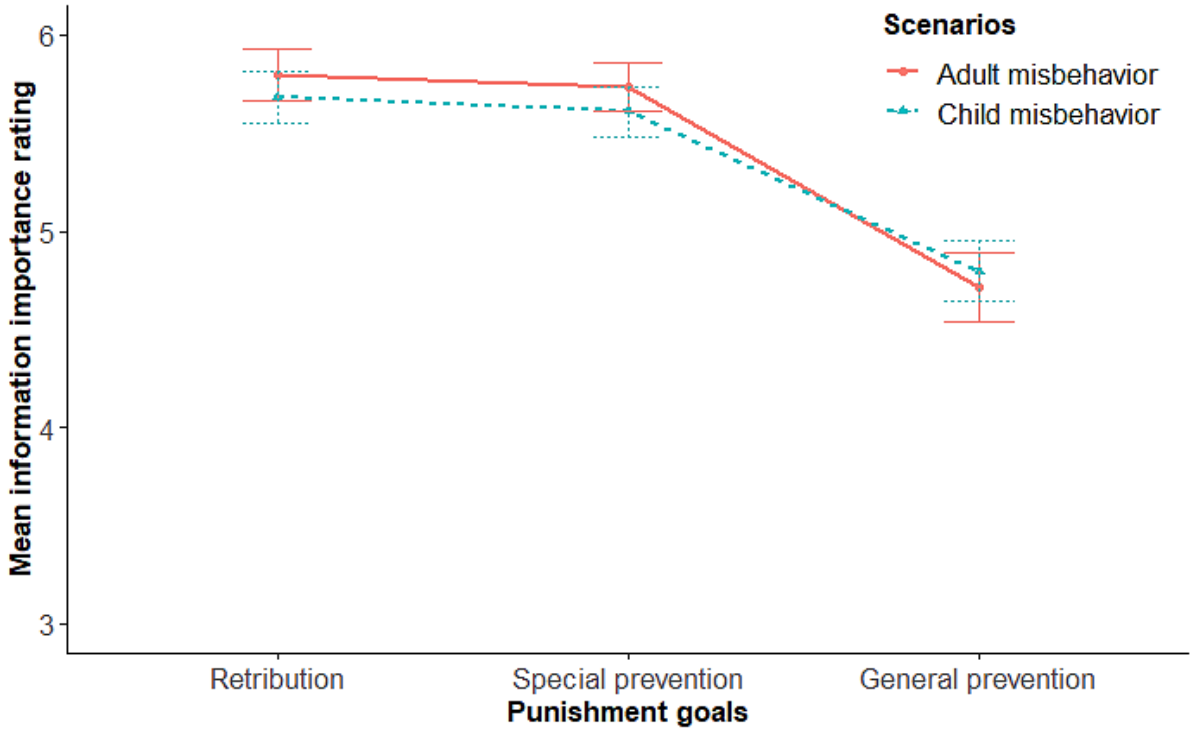

Figure 2. Mean information importance ratings of the three punishment goals in each scenario. Error bars represent one standard error of the mean.

In contrast to the results of the information-selection task (i.e., both scenarios), post-hoc *t*-tests revealed no significant differences between retribution and special prevention as punishment goals, $t(68) = 0.60$, $p = .548$, $d = 0.07$, 95% CI [-0.26, 0.41]. However, items associated with retribution were rated as significantly more important than items associated with general prevention, $t(68) = 6.75$, $p < .001$, $d = 0.81$, 95% CI [0.46, 1.16]. Likewise, items associated with special prevention were rated as significantly more important than items associated with general prevention, $t(68) = 6.49$, $p < .001$, $d = 0.78$, 95% CI [0.43, 1.13].

Overall, Study 1 revealed retribution to be the most important punishment goal of children, special prevention was the second important punishment goal, and general prevention was the least important punishment goal. These general patterns were consistent across both cases of misbehavior, that is, it was independent of whether misbehavior was committed by a child

or an adult. However, results were somewhat inconsistent concerning the relative importance of special prevention: once forcing children to choose between punishment goals (or information associated with these goals) in the information-selection task, special prevention received significantly lower scores than retribution. By contrast, once allowing children to indicate the importance of each of the three punishment goals separately, special prevention was rated as almost equally important as retribution.

In summary, although we found retribution to be the most important punishment goal of children (whereas general prevention was least important), this study revealed partially mixed results regarding the role of special prevention. Specifically, when punishment goals were in competition with each other due to the forced choice type of method applied in the information-selection task (i.e., the support of one punishment goal leads to the rejection of the other punishment goals), retribution was clearly superior over special prevention and general prevention. However, when children were able to indicate the importance of each punishment goal separately (i.e., the support of one punishment goal does not affect the support of the other goals), the dominance of retribution as the driver of punishment behavior diminished—special prevention received almost similar importance ratings. However, given that the information importance rating task revealed considerable ceiling effects, we cannot interpret these diminished differences between retribution and special prevention conclusively. In light of these mixed results, we considered it vital to further investigate children's punishment goals in a second study.

3.4 STUDY 2

In this second study, we extended Study 1 in primarily two aspects. First, given that punishment most often serves multiple goals at the same time (Gromet & Darley, 2009a), forcing children to choose between these goals appears rather artificial. Therefore, we again applied

a methodological approach assessing the importance of each punishment goal in children independently. Second, in Study 1 we investigated children's motives underlying uncostly punishment. However, these need not match the goals children pursue in a situation of costly third-party punishment. Therefore, we adapted an economic game recently introduced to study the punishment goals of adults and modified it to (i.) make it comprehensible for children and (ii.) to differentiate between special prevention and general prevention as two separate goals of punishment (Crockett et al., 2014). In this paradigm, participants observe another person accumulating own payoffs by acting unfairly. Participants then have the opportunity to reduce this person's payoff by sacrificing own resources. To be able to isolate the unique effects of different punishment motives, it is controlled whether the offender (special prevention), a bystander (general prevention), or nobody (retribution) learns about the punishment across various experimental conditions. Based on the prior findings, we expected retribution to be of particular importance for children's punishment, whereas general prevention should be least important. Given the mixed results regarding special prevention, we had no clear hypotheses about the role of this goal for children's punishment behavior.

3.4.1 Methods

Measures and procedure

Data collection took place in three German public schools. On arrival in the classroom, children provided the experimenters with the consent form signed by their parents or legal guardians. Then, children were seated individually in front of computers. Two experimenters welcomed the class and gave them comprehensive verbal instructions on the subsequent task, which was a third-party punishment game. Children (in the role of the punisher) were asked to imagine working on a task with three other children: Child A (i.e., the offender), Child B (i.e., the victim), and Child C (i.e., the bystander). They were told that they do not know these other

children, nor will they ever meet them as the task takes place in a virtual chatroom.[2] Children learned that they were going to work on four rounds of this task, with changing individuals in the other roles.

The rules of each round were as follows: initially, the punisher received five piggy banks and the offender received 10 piggy banks. These piggy banks contained an undefined number of popular coins. In the first step, the offender decided how many of her own piggy banks should be shared with the victim. In the second step, the punisher privately opened all piggy banks and counted the number of coins each child (i.e., the punisher, the offender, and the victim) had. Subsequently, the punisher had the opportunity to reduce the offender's payoff by spending own coins; each coin spent by the punisher resulted in a reduction of two coins for the offender. After the punisher's decision, the round was over and all children in the chatroom received their final payoff. The victim and the bystander remained inactive throughout.

To be able to isolate the unique effects of punishment motives (retribution vs. special prevention vs. general prevention), we manipulated the offender's and the bystander's knowledge of whether the punisher decided to reduce the offender's payoff across four experimental conditions. Although the offender's payoff was always reduced when the punisher spent coins to do so (and the punisher knew this), whether the offender or the bystander learned about the decision of the punisher varied across conditions. In the *hidden punishment condition*, neither the offender nor the bystander was informed about whether the punisher had spent own coins to reduce the offender's payoff. Given that punishment was not communicated to the offender or the bystander in this condition, it could not prevent future misbehavior and, therefore, punishment in this condition can be described as purely retributive (i.e., it is not motivated

[2] There was also a fifth person in the chatroom that we labeled "group manager." This person's function was merely to structure the interaction in the chatroom, reminding the participating children of the procedure of the task.

by preventing future misbehavior). In the *private punishment condition*, the offender was informed whether the punisher had reduced the offender's payoff as a reaction to the decision made in step 1. Thus, the offender learned whether the own behavior shown in step 1 is legitimate and goes unpunished, or not (i.e., adding special prevention as a punishment motive to retribution). In the *public punishment condition*, the bystander was informed whether the punisher had spent own coins to reduce the offender's payoff as a reaction to the decision of the offender made in step 1. Thus, the bystander learned whether the offender's behavior shown in step 1 is legitimate and goes unpunished, or not (i.e., adding general prevention as a punishment motive to retribution). In the fourth condition, actual punishment motives were separated from other payoff-based motives (e.g., inequality aversion or spite), as these have been shown to largely affect children's behavior (e.g., Hamann et al., 2011; Lobue et al., 2011; McAuliffe, Blake, & Warneken, 2014; Schmidt & Sommerville, 2011). In this *control condition*, the offender acted fairly towards the victim but still ended up with more coins than the victim due to luck (i.e., the offender's advantage in coins was unintentional). Therefore, the offender's behavior should not be interpreted as an offense requiring a punishment reaction by the punisher, but the punisher could take a decision to reinforce equality. Punishment in this condition was neither communicated to the offender nor to the bystander and can thus be ascribed to purely payoff-based motives.

To reduce complexity for the children, they only participated in the game in the role of the punisher. Thus, the interactions in the chatroom were only hypothetical. Children were informed about this because deception is not aligned with the ethical principles of psychologists, is very rarely justifiable (Hertwig & Ortmann, 2008), and has wide-ranging costs for the future perception of psychological research by participants (Ortmann & Hertwig, 2002). However, children were urged to imagine as best as they can that this is a real interaction. To further reduce complexity, the differences between conditions were explicitly explained to the partici-

pating children, making sure that they knew that, for example, punishment in the hidden condition could not be used to teach the offender a lesson. Moreover, all participating children started with the control condition, followed by the hidden condition (i.e., two rounds with the same rules but different behavior of the offender). After explaining the changes in the rules (i.e., other children in the chatroom will learn about the punisher's decision), children worked on the private and, lastly, on the public condition. Whereas, the offender divided the piggy banks fairly in the control condition (i.e., giving half of the 10 piggy banks to the victim), the offender acted very unfairly in all other conditions (i.e., giving only two of the 10 piggy banks to the victim). Lastly, to once more reduce complexity for the children, the distribution of coins was consistent across the four conditions (i.e., the offender had 24 coins, the victim had six coins, and the punisher had 10 coins). Thus, children were always provided with 10 coins to reduce the offender's payoff.

Following the third-party punishment game, children provided demographic information and then reached a stop sign asking them to wait until everyone else had finished this part of the study. Subsequently, they worked on an unrelated task, before they were thanked and debriefed.

Sample

Approximating power for hierarchical linear models using simulations requires concrete values for several parameter estimates in the model (Bolger & Laurenceau, 2013), which needs previous data to be based on. Thus, as this is the first study using a new paradigm (and therefore previous data was not available), we were not able to determine sample size using simulations. Rather, we used more straightforward sample size calculations, conducting an a priori power analysis for an ordinary repeated measures ANOVA using G*Power (Faul et al., 2009, 2007). Specifically, we aimed to detect a small to medium-sized effect of $f = .15$ in a repeated measures ANOVA (number of groups = 1, number of measurements = 4 for the four conditions) with

high power of 1-β = .90, given a conventional α = .05, and nonsphericity correction ε = 1. This resulted in a required sample size of N = 213 participants. Furthermore, comparable to Study 1, a pragmatic approach was utilized before starting data collection, as we again collaborated with schools for this project (und thus data collection could not simply be interrupted once reaching the required sample size). Therefore, we planned to collect data in three schools and to recruit further schools for participation if we did not reach the threshold of approximately 215 participating children.

We collected data in the fifth and sixth grade of three public German schools. After 12 sessions, N = 238 children had participated in the study. Around 45% of participants (i.e., n = 106) were female, most children's mother language was German (92%), and ages ranged between 9 and 12 (M = 10.46, SD = 0.61; one child did not indicate any age).

3.4.2 Results and discussion

Figure 3 displays the coins spent for punishment in the four conditions across participants. As shown, the mean number of coins spent to reduce the offender's payoff was lowest in the control condition (M = 4.55, SD = 2.65), although children already indicated a notable degree of payoff-based motives (i.e., spite or inequality aversion), as they sacrificed almost half of their own credit in this condition. These generally high levels of punishment can also be ascribed to the widely demonstrated intrinsic human aversion to inequity (Dawes, Fowler, Johnson, McElreath, & Smirnov, 2007), which has also been extensively observed in children (e.g., Hamann et al., 2011; Lobue et al., 2011; McAuliffe et al., 2014; Schmidt & Sommerville, 2011). Punishment was slightly stronger in the hidden punishment condition (M = 5.10, SD = 2.82) and clearly strongest in the two preventive conditions private punishment (M = 6.18, SD = 2.69) and public punishment (M = 6.38, SD = 2.64), which turned out roughly comparable.

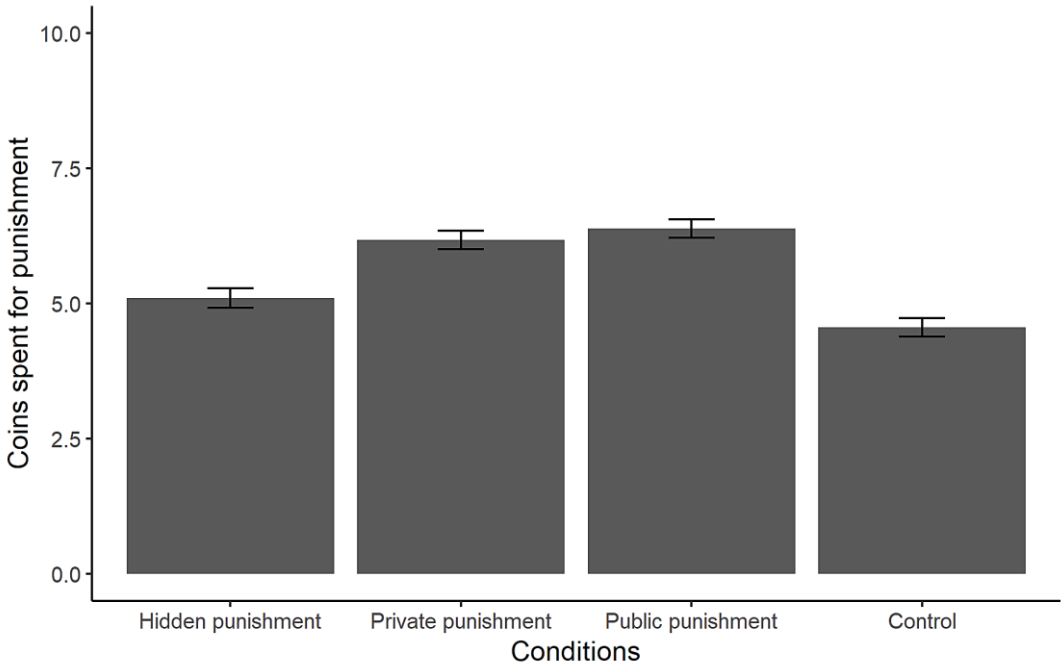

Figure 3. Coins spent for punishment in the four conditions across participants. Error bars represent one standard error of the mean.

To test the specific impact of different punishment motives on children's punishment behavior, the four experimental conditions were contrast-coded into three dummy variables (D1, D2, D3). D1 reflects the contrast for retribution, comparing the control condition (i.e., fair behavior of the offender and an advantage in coins due to luck) with the hidden condition (i.e., unfair behavior of the offender with hidden punishment). A significant effect of D1 indicates a punishment beyond an intervention due to purely payoff-based motives such as inequality aversion and without the potential benefit of preventing future misbehavior. Thus, punishment in this condition is driven by retribution. D2 reflects the contrast for special prevention, comparing the hidden condition (i.e., unfair behavior of the offender with hidden punishment) with the private condition (i.e., unfair behavior of the offender with a punishment communicated to the offender). A significant effect of D2 indicates a punishment beyond an intervention due to retributive purposes and is therefore motivated by special prevention (i.e., teaching the offender a lesson). D3 reflects the contrast for general prevention, comparing the hidden condition (i.e., unfair behavior of the offender with hidden punishment) with the public condition (i.e., unfair

behavior of the offender with a punishment communicated to the bystander). A significant effect of D3 indicates a punishment beyond an intervention due to retributive purposes and is therefore motivated by general prevention (i.e., communicating to the bystander that unfair behavior cannot go unpunished). All dummy variables were entered simultaneously into a hierarchical linear model. No assumptions regarding the structure of the variance-covariance matrix of the three dummy variables were made.

The general differences between conditions (payoff-based reaction vs. hidden punishment vs. communication to the offender vs. communication to the bystander) had a significant effect on punishment, $\chi^2(3) = 116.42$, $p < .001$. Contrasts testing the particular impact of each punishment goal on children's punishment behavior revealed a significant effect of D1 that reflects the expected differences between fair and unfair behavior with a non-communicated punishment, and thus, indicates a retributive punishment beyond an intervention due to purely payoff-based motives such as inequality aversion, $b = 1.00$, $t(711.00) = 8.58$, $p < .001$. Furthermore, there was a significant effect of D2, reflecting greater punishment in the condition allowing the punisher to teach the offender a lesson than in the condition in which punishment was purely retributive, $b = 0.62$, $t(711.00) = 5.35$, $p < .001$. Thus, children's punishment was also largely motivated by special prevention, which is in line with the results of the information importance rating in Study 1. Surprisingly, D3 was also significant, reflecting greater punishment in the condition allowing the punisher to communicate a norm of fairness (in which misbehavior gets punished) to a bystander than in the condition in which punishment was purely retributive, $b = 0.83$, $t(711.00) = 7.12$, $p < .001$. Therefore, children's punishment was also largely motivated by general prevention, which we did not expect given the results of Study 1.

In sum, Study 2 replicated Study 1 in terms of the importance of retribution as a punishment goal of children. Additionally, we found further evidence for special prevention as an important motive underlying children's punishment behavior. This is in line with the result of the information importance rating task applied in Study 1, in which punishment goals were also

not "artificially" juxtaposed, but evaluated separately. However, Study 2 revealed differences to Study 1 regarding the importance of general prevention as a punishment goal of children. Whereas general prevention was consistently evaluated as the least important punishment goal in Study 1, it received substantial support in the economic game approach applied in Study 2, identifying general prevention as a punishment goal of children that is of comparable importance to retribution and special prevention.

3.5 GENERAL DISCUSSION

There is considerable research showing that already young children engage in punishment when they themselves have been harmed (Gummerum & Chu, 2014; Robbins & Rochat, 2011), are willing to intervene and correct behavior that harms others (Hamlin et al., 2011; Kenward & Östh, 2012; Riedl et al., 2015), and sacrifice own resources for third-party punishment when others are the victims of norm violations (Jordan et al., 2014; McAuliffe et al., 2015). However, the motives underlying such third-party punishment of children has, so far, not received much attention. Therefore, we conducted two studies investigating the extent to which children's punishment behavior is driven by retribution (i.e., to even out the harm caused), special prevention (i.e., to prevent future misconduct by the offender), and general prevention (i.e., to prevent future misconduct by others; Carlsmith, 2006; Carlsmith et al., 2002). We adapted several methodological approaches to measure children's punishment goals in ways that do not require children to introspectively predict the goals they might intend to achieve when punishing an offender, as such approaches have been shown to be of rather questionable validity for actual punishment behavior (e.g., Crockett et al., 2014).

In Study 1, children were asked to determine a punishment for two scenarios describing norm violations. Before expressing their punishment, they were provided with six pieces of information on the misbehavior that are known to be associated with either of the three punish-

ment goals (Keller et al., 2010). Then, children were asked to (i.) select three of these information pieces (in an information-selection task) and, in a second task, (ii.) rate the importance of these pieces of information for their punishment decision (in an information importance rating task). Across both measures, retribution was consistently revealed as the most important punishment goal of children, with special prevention as the second and general prevention as the least important punishment goal.

However, whereas retribution was significantly more important than special prevention in the information-selection task, these differences largely disappeared in the information importance rating task. Importantly, the diminished differences between retribution and special prevention may have been due to the ceiling effects present in the information importance rating task. Nonetheless, the inconsistencies between methods may have also been caused by methodological differences, as the information-selection task forces children to choose between punishment goals (or at least the information associated with these goals) rather than measuring support of punishment goals independently. Arguably, such an approach appears somewhat artificial, given that punishment behavior typically serves multiple goals simultaneously (Gromet & Darley, 2009a).

The impact of special prevention on children's punishment decisions was also revealed in Study 2, in which we adapted a third-party punishment game that has recently been introduced to examine people's punishment goals (Crockett et al., 2014). Further deepening our understanding of the motives underlying actual altruistic punishment of children, punishment in this economic game was introduced as being costly. In the game, children observed another child accumulating payoffs by acting unfairly. Children then had the opportunity to reduce this child's payoff by sacrificing own resources. To be able to isolate the unique effects of the punishment motives, it was controlled whether the offender (i.e., important for special prevention), a bystander (i.e., important for general prevention), or nobody (i.e., retribution) learned about

the punishment across various experimental conditions, with the idea that punishment can only serve preventive goals if others are aware of it.

This altruistic third-party punishment paradigm supported the patterns found in Study 1, with retribution as well as special prevention being essential motives underlying children's punishment behavior. In contrast to Study 1, general prevention appeared to be comparably important to retribution and special prevention. That is, compared to a hidden punishment situation (i.e., nobody learns about the punishment), children were not only willing to sacrifice more resources when the offender learned about the punishment (i.e., special prevention), but also when a bystander learned about the punishment (i.e., general prevention). The differential impact of general prevention in Study 1 and Study 2 may be due to several differences between the studies, such as the different types of misbehavior to be punished or whether punishment was uncostly or costly. However, as we can only speculate about what differences of the two studies changed the role of general prevention in children's punishment, further research is needed to identify the conditions under which general prevention gains children's support.

The results of the present research have several implications, both theoretical and practical. On a theoretical level, the findings add to an increasing body of research shedding doubts on the idea of people as "intuitive retributivists" (e.g., Goodwin & Benforado, 2015). Both our studies challenge the "intuitive retributivists" position, because third-party punishment in children was also largely motivated by utilitarian purposes (at least by special prevention). Moreover, the present studies again provide evidence that the results of research on punishment goals are subject to the methodological approaches applied, further raising awareness of potentially underestimating punishment goals that actually drive punishment behavior, mainly due to methodological considerations (Goodwin & Benforado, 2015).

On a practical level, the results are important both for teachers and parents handling children's misbehavior on a daily basis (Kulinna et al., 2006). In school, student misbehavior is a main cause of conflicts, regularly leading to feelings of injustice in students (Fan & Chan,

1999; Israelashvili, 1997). Knowing that both retribution and special prevention are goals that are generally supported by children could provide teachers with more confidence in the treatment of child misbehavior, something many teachers are aspiring (Melnick & Meister, 2008). However, in light of the rather mixed results on the role of general prevention, practitioners should be cautious about applying punishment that appear to primarily pursue the prevention of imitation, as differences between the expected and actually pursued punishment goals of authorities may lead to a decreasing compliance to rules and norms by the children (Mooijman et al., 2017).

Of course, there are potential limitations of the present studies that are worth being discussed. Most importantly, although we applied a third-party punishment game in Study 2, children did not actually interact with each other and thus, did not make real consequential decisions. Given the importance of ethical consideration in psychological sciences (Hertwig & Ortmann, 2008; Ortmann & Hertwig, 2002), we refrained from deceiving the participating children but relied on hypothetical interactions. However, we encouraged children to imagine that the decisions made in the chatroom are part of a truly consequential interaction.

Indeed, recent research has shown that differences between decisions in actual vs. hypothetical economic games are remarkably small (e.g., Ben-Ner, Amit, & Levy, 2008; Engel, 2011), and that the impact of social desirability on behavior in such games is, to a great extent, rather subject to the perceived anonymity (Thielmann, Heck, & Hilbig, 2016). Thus, we attached great importance to create an atmosphere of anonymity when conducting this research in the classes (e.g., by restructuring the classroom). Furthermore, we have no reason to assume any differences between conditions regarding the impact of social desirability. Thus, if children may have been biased to hypothetically punish more than they would probably do in a truly incentivized situation, this should be equally true for all conditions and should, therefore, have no impact on the results derived from within-subjects comparisons. Hence, we have sufficient

confidence that potential biases, at most, were consistent across conditions and therefore should not have had an impact on the results and conclusions derived.

In sum, the present research is the first to investigate the motives underlying third-party punishment in children. Across two studies relying on different methodological approaches, we found that children's punishment is not only motivated by retribution (i.e., to even out the harm caused), but also by utilitarian punishment goals—certainly by special prevention (i.e., to prevent future misconduct by the offender) and, at least in some cases, also by general prevention (i.e., to prevent future misconduct by others). These results shed initial light on children's punishment goals and thus, both contribute to an ongoing scientific discussion on the primary purposes of third-party punishment (Goodwin & Benforado, 2015; Goodwin & Gromet, 2014) and provide practical implications for the treatment of child misbehavior (e.g., in school; Melnick & Meister, 2008).

4. PROJECT 2: "PUNISHMENT GOALS IN CLASSROOM INTERVENTIONS: AN ATTRIBUTIONAL APPROACH"

4.1 ABSTRACT

Individuals' punishment goals depend on the perceived cause of the misbehavior. However, a corresponding attributional model of punishment goals has only been studied in legal domains—but was largely ignored in others, such as the educational domain, in which student misbehavior and its treatment is a main stressor for both teachers and students. Thus, we investigated teachers' punishment goals in classroom settings depending on their attribution of student misbehavior. Specifically, we asked laypeople (Experiment 1, $N = 233$), pre-service teachers (Experiment 2, $N = 119$), and in-service teachers (Experiment 3, $N = 141$) to read several versions of a scenario describing a student destroying the belongings of another student. Using a 2 x 2 within-subjects design, we manipulated the stability (stable vs. unstable) and controllability (controllable vs. uncontrollable) of the cause of the misbehavior. Results show that the support of retribution (i.e., evening out the harm caused) as a punishment goal in classroom interventions is largely independent of the perceived cause of the misbehavior. By contrast, the support of special prevention (i.e., preventing future misbehavior by the offending student) and general prevention (i.e., preventing future misbehavior by other students) is primarily subject to the perceived controllability of the misbehavior. Overall, we show that models of punishment behavior developed in other domains cannot simply be applied to teachers' classroom intervention strategies.

Punishment goals in classroom interventions: An attributional approach

In October 2012, Lance Armstrong—a former professional road racing cyclist and one of the most successful athletes in history—was banned retrospectively by the International Cycling Union due to long-term doping offenses. Correspondingly, all results dating back until 1998 were voided and he was excluded from Olympic sports for life. Strikingly, this was realized although Armstrong had already announced his retirement from competitive cycling, thus raising the question about the goal of this act of punishment given that future transgressions by Armstrong were unlikely, not to say impossible: Was it in order to restore justice for the competitors losing against Armstrong? Or was it to set an example that doping in (cycling) sports will not pay off, but cheaters will eventually be caught and brought to justice?

According to social and behavioral sciences, the threat of punishment is a key factor to enforce adherence to norms and laws in a world of incentivized selfishness, both on a macro-level (e.g., entire countries and societies) and a more micro-level (e.g., in interpersonal contacts or institutions; Boyd et al., 2010, 2003; Fehr & Gächter, 2002). This perspective reflects a regulatory function of punishment, based on the idea that the fear and experience of being punished will arguably lead to a more cooperative society and less future crimes (Henrich et al., 2006). Following such a *utilitarian* approach, the primary goal of punishment is to prevent future misbehavior by the perpetrator herself (termed *special prevention*) or by unrelated observers (termed *general prevention*; Keller et al., 2010; Rucker et al., 2004).

A normatively different approach of punishment, in turn, suggests a rather backward-oriented goal of punishment, namely, paying back harm doers for their misbehavior (Goodwin & Gromet, 2014; Keller et al., 2010). This approach is referred to as *retribution*, where the primary goal of punishing offenders is to rebalance the moral wrong that has been committed by the offense and, thus, to restore justice. Correspondingly, retribution is sometimes also referred to as “the just deserts” theory (Carlsmith et al., 2002). Indeed, research suggests that

punishment is, overall, primarily driven by retribution (i.e., to even out the wrong that has been done) rather than special or general prevention (i.e., to prevent future crimes by the offender or others; Carlsmith, 2006; Carlsmith et al., 2002).

However, people may indeed adjust their punishment goals depending on aspects of the situation, such as the punisher's status in the situation or the circumstances in which the misbehavior occurred. For example, it has been found that individuals in a more powerful position (e.g., managers) show a greater preference for utilitarian punishment goals (Mooijman et al., 2015). Furthermore, and most importantly for the present work, research suggests that the perceived cause of the misbehavior is a fundamental factor for individuals' punishment goals. This is in line with an attributional account of punishment (Graham et al., 1997; Weiner et al., 1997).

According to attribution theory, behavior can be attributed on three dimensions: its locus or location (internal or external to the actor), its stability (whether the cause is persistent across time or temporary), and its controllability (whether the cause was subject to the actor's volition or not). With regard to the decision to punish an offender, prior research has shown that especially the perceived stability and controllability of the cause of the misbehavior influence the goals individuals intend to achieve (Graham et al., 1997; Weiner et al., 1997). The attributional model of punishment goals (Graham et al., 1997; Weiner et al., 1997) suggests that, if the cause of misbehavior is perceived as stable, one anticipates future misbehavior by the perpetrator and therefore assumes that the efficacy of interventions is low, leading to less focus on utility in punishment. By contrast, if the cause of misbehavior is perceived as unstable, future misbehavior by the perpetrator is less predictable, implying that interventions may be more effective, in turn leading to a stronger focus on utility in punishment. Likewise, if the cause of misbehavior is perceived as controllable, the perpetrator is credited with more responsibility; therefore, anger increases and sympathy decreases, leading to a stronger focus on retributive punishment.

By contrast, if the cause of misbehavior is perceived as uncontrollable, the perpetrator is credited with less responsibility, anger decreases and sympathy increases, leading to a lower focus on retributive punishment.

In the present research, we apply this attributional view on punishment goals in a specific context in which misbehavior and its punishment happen on a daily basis but are nonetheless very poorly understood: the school environment. Teachers frequently have to deal with a wide range of student misbehaviors (e.g., Kulinna et al., 2006; Wheldall & Merrett, 1988). At the same time, teachers report a high extent of insecurity in such situations (Melnick & Meister, 2008). In turn, it appears that trying to find a just solution in the case of student misbehavior regularly ends up in conflicts, causing feelings of injustice in students (Fan & Chan, 1999; Israelashvili, 1997). A plausible aspect of teachers' punishment behavior that may cause such conflicts are the goals teachers intend to achieve when reacting to student misbehavior. The present research aims to illuminate this issue by applying an attributional approach to understand and study teachers' punishment behavior, and in particular, the goals they intend to achieve when reacting to misbehavior.

To the best of our knowledge, an attributional analysis of teachers' behavior in the classroom has only once been applied in prior research (Reyna & Weiner, 2001). However, in these studies, the authors focused on the reaction of teachers to students failing an exam rather than engaging in misbehavior. Thus, the studies measured the goal of providing feedback to underachieving students rather than the goal of punishing students and, correspondingly, only considered special prevention as a potential goal of a teacher's reaction. This focus on special prevention may be sufficient in the case of providing feedback, as performance in exams is typically given in private. However, classroom misbehavior usually occurs publicly: it is observed by—and may even directly affect—other students. Thus, preventing imitation by these observers (i.e., general prevention) should arguably be an important goal of teachers when reacting to

misbehavior. Therefore, it is vital to consider general prevention when studying teachers' punishment goals.

Additionally, participants in Reyna and Weiner's (2001) studies were asked to indicate the endorsement of different "punishment goals" they intended to achieve when giving their feedback. However, there is a growing body of research suggesting a rather weak correlation between individuals' endorsement of punishment goals and their actual punishment behavior. Specifically, it has been shown that although people highly endorse *all* goals of punishment and, additionally, indicate a substantial endorsement of utilitarian punishment goals, their actual punishment *behavior* is mostly retributive (Carlsmith, 2008; Crockett et al., 2014). Thus, relatively "direct" measures of punishment goals—explicitly stating a punishment goal and asking participants for their endorsement—may be biased as individuals overestimate the influence of utilitarian goals on their own punishment decisions. Consequently, it is important to consider additional, less direct or explicit measures of punishment goals. Such a more indirect measure of people's punishment goals is to present participants with specific punishment reactions (which may serve certain goals without stating these explicitly) and to assess the perceived appropriateness of these reactions. To the extent that the reactions are known (e.g., from pretests) to differ on how well they serve different punishment goals—that is, retribution, special prevention, and general prevention—one can consider such appropriateness ratings as indirect measure of punishment goals that does not require individuals to introspectively predict the exact goals they might intend to achieve.

4.2 THE PRESENT RESEARCH

The goal of the present work is to investigate teachers' punishment goals in classroom contexts. As sketched above, prior research suggests that individuals' punishment behavior is primarily driven by retribution (Carlsmith et al., 2002), although people are able to adjust their punishment goals depending on the perceived cause of the misbehavior, as predicted by an

attributional approach (Graham et al., 1997; Weiner et al., 1997). Herein, we apply such an approach to provide insights on teachers' treatment of student misbehavior, specifically the goals they intend to achieve. In three experiments, we (i.) experimentally manipulated the attribution of student misbehavior in terms of the stability and controllability of its cause, (ii.) differentiated between special and general prevention as two separate utilitarian goals of teachers' classroom intervention strategies, and (iii.) used both a "direct" measure of goal-endorsement and an "indirect" measure of punishment goals based on appropriateness ratings of different, pretested reactions to a student's misbehavior. We collected data from individuals without any pedagogical background as well as from pre-service and in-service teachers to potentially identify the effects of teaching experiences on attribution of and intervention following student misbehavior (e.g., Andreou & Rapti, 2010).

Based on the attributional analysis of punishment goals in other domains (Graham et al., 1997; Weiner et al., 1997) as well as on prior research on teachers' reactions to students failing an exam (Reyna & Weiner, 2001), we hypothesized an effect of the perceived stability of the cause of the misbehavior on the support of utilitarian (i.e., special preventive) punishment, with attribution of the misbehavior to a stable cause leading to a lower support than attribution to an unstable cause. Further, we expected that the perceived controllability of the cause of the misbehavior should have an influence on the support of retributive punishment, with attribution to a controllable cause leading to a higher support than attribution to an uncontrollable cause. However, given that there are major differences between prior research and our experiments—and because there is, to the best of our knowledge, no research examining the effect of causal attribution on the endorsement of general prevention—we consider our first experiment as rather exploratory. In Experiment 2 and 3, we then seek to replicate all results from Experiment 1 with samples with a larger pedagogical background.

4.3 EXPERIMENT 1

In the first experiment, we aimed to collect initial evidence on the influence of the attribution of a student's misbehavior on teachers' punishment goals. Given that these data sought to provide a first exploratory test, we did not seek a sample with pedagogical background, but collected data from a convenience sample. All materials, the data, and analyses (including outputs) of all three experiments are available online at the Open Science Framework (OSF; accessible during double-blind peer-review via the following link: https://osf.io/fvj3a/?view_only=ec95615330c84efe835e00b95a9d70d2).

4.3.1 Methods

The experiment implemented a 2 x 2 within-subjects design. Specifically, participants read several versions of a scenario describing a student destroying the belongings of another student and manipulated the stability (stable vs. unstable) and controllability (controllable vs. uncontrollable) of the cause of the misbehavior. Participants were asked to indicate their perceptions of and reactions to these scenarios. As sketched above, we measured participants' punishment goals both directly (i.e., asking for the endorsement of punishment goals) and indirectly (i.e., asking for the appropriateness of concrete punishment reactions). The material to measure punishment goals indirectly via appropriateness ratings was pretested in a separate study as outlined below.

Pretest of experimental material: Test of reactions

To measure participants' punishment goals indirectly, we created several reactions to a scenario describing the misbehavior of a student. This misbehavior was the same as in the main experiments, that is, a student destroying a recently prepared hand drum of another student. Participants (N = 122) were asked to rate nine reactions with regard to the degree to which they served each of the following punishment goals: retribution, special prevention, and general prevention. For the main experiment, we aimed to identify three reactions, each of which would

primarily serve one of these three punishment goals but not the other two. Additionally, we considered it important that the reactions were perceived as equally severe. All materials, the data, and detailed results (including test statistics) of this pretest are available online at the OSF. The three selected reactions that were used in the main experiments are provided in Table 1.

Table 1

Punishment reactions derived from the pretest and used in the main experiments

Primary punishment goal	Reaction
Retribution	You enter X's misbehavior—without saying a word—to the class register[3]. After the lesson, you tell X to repair Y's drum.
Special Prevention	After the lesson, you tell X to write an essay of three pages at home about why things of others should not be damaged.
General prevention	You tell the whole class that you will enter X's misbehavior to the class register. In case of future rule-breaking by any student, you will also inform the parents.

Measures and procedure

The main experiment was conducted via the Internet. After providing informed consent, participants were asked to imagine being a teacher of a class whose students they already knew. Participants next received five scenarios describing a student destroying a recently prepared

[3] In Germany, a class register is a notebook in which teachers make daily notes about all important aspects of their lessons (e.g., the topic of the lessons, absence of students) and particular incidents that happened in class (e.g., student misbehavior). The school principal reviews all class registers regularly and takes action if she disapproves with something (e.g., if a student is misbehaving frequently). Furthermore, the class register may be considered for grading students at the end of the year.

hand drum of another student. The first scenario was neutral in that it did not include any experimental manipulation of the attribution of the cause of the misbehavior[4]; the subsequent four scenarios, in turn, were manipulated with regard to the two causal attribution dimensions of stability (S) and controllability (C). All causes were internal to the student, that is, the student was responsible for breaking the hand drum. The four scenarios including a manipulation were presented in randomized order. Each of them consisted of the same opening statement ("Within the past lessons, you manufactured hand drums with your students that you plan to use today. You briefly turn to the board. Once you face the class again,…"), followed by the manipulated information as follows:

> Stable & Controllable (S/C) scenario: "…you see that Florin **deliberately punctures Maxi's drum with a pencil**. As a result, the drum is broken. Florin is the **troublemaker of the class** and regularly runs riot. It is **already the fourth time** this year that Florin has destroyed the belongings of a class member."
>
> Stable & Uncontrollable (S/UC) scenario: "…you see **Florin fooling around** and as a result Maxi's drum falls to the ground. The fall punctures a hole and thus the drum is broken. Florin is **very clumsy** and regularly puts his foot in his mouth. It is **already the fourth time** this year that Florin has destroyed the belongings of a class member."
>
> Unstable & Controllable (US/C) scenario: "…you see that Florin **deliberately punctures Maxi's drum with a pencil**. As a result, the drum is broken. **Usually**, Florin is rather **inconspicuous**. This is the **first time** that Florin has destroyed the belongings of a class member."

[4] As this scenario was included for a different, unrelated research question, it will not be further considered in the present article and also not included in any analysis reported in what follows. Nonetheless, the corresponding material and data are available at the OSF.

> Unstable & Uncontrollable (US/UC) scenario: "…you see **Florin fooling around** and as a result Maxi's drum falls to the ground. The fall punctures a hole and thus the drum is broken. **Usually**, Florin is rather **inconspicuous**. This is the **first time** that Florin has destroyed the belongings of a class member."

Following each scenario, participants were asked to rate the *stability* and *controllability* of the cause of the student's misbehavior.[5] Next, they completed the *indirect measure of punishment goals*. That is, we asked participants to rate the appropriateness of each of the three reactions extracted from our pre-study (see above) using a 6-point Likert-type scale ranging from 0 = "*not at all appropriate*" to 5 = "*completely appropriate.*"

Finally, participants completed a *direct measure of punishment goals*, indicating the goals they would want to accomplish if presented a chance to react to the misbehaving student. To this end, we adapted one item for each punishment goal (retribution, special prevention, and general prevention) from Orth (2003) and Weiner and colleagues (1997). Each item was answered on a 6-point scale ranging from 0 = "*not at all*" to 5 = "*completely*," with higher values indicating stronger endorsement of a particular punishment goal. After completing all scenarios, participants provided demographic information, before they were fully debriefed and thanked.

[5] Additionally, participants were asked to indicate the student's *responsibility* for what occurred, how much *anger* and *sympathy* they would feel towards the misbehaving student as well as to what degree they thought their own behaviors (e.g., reinforcements, feedback, and discipline) could *influence* the student's future behavior. These data were collected following Reyna and Weiner (2001). However, because it is not of primary interest for our hypotheses, correlations of these measures are exclusively reported in the supplemental material at the OSF.

Sample

To determine the required sample size, we conducted an a priori power analysis using G*Power (Faul et al., 2009, 2007). Specifically, we aimed to detect a small to medium-sized effect of $f = .15$ in a repeated measures ANOVA (number of groups = 1, number of measurements = 4 for the four manipulated scenarios) with high power of $1\text{-}\beta = .90$, given a conventional $\alpha = .05$, and nonsphericity correction $\varepsilon = 1$. This resulted in a required sample size of $N = 213$ participants.

Participants were recruited through mailing lists, social media platforms, and personal contacts, and invited to take part in an experiment lasting around 15 minutes. As an incentive for participation, participants had a chance to win one of twenty 20€ gift vouchers. Overall, 328 participants started the experiment, of which 233 (71%) completed it. The majority of participants were female (i.e., $n = 172$; 74%). Participants' ages ranged between 18 and 58 years ($M = 26.96$, $SD = 8.26$) and most were students (48.07%) or employees (32.19%).

4.3.2 Results and discussion

Manipulation check

We found the manipulation of both the stability and controllability of the cause of the student's misbehavior to be successful. The two scenarios representing stable causes received higher ratings of stability ($M_{S/C} = 4.08$, $SD_{S/C} = 1.06$, and $M_{S/UC} = 3.45$, $SD_{S/UC} = 1.22$, respectively) compared to the two scenarios with unstable causes of the misbehavior ($M_{US/C} = 1.80$, $SD_{US/C} = 1.33$, and $M_{US/UC} = 1.30$, $SD_{US/UC} = 1.25$, respectively). Likewise, the two scenarios with controllable causes of the misbehavior received higher ratings of controllability ($M_{S/C} = 3.48$, $SD_{S/C} = 1.34$, and $M_{US/C} = 3.71$, $SD_{US/C} = 1.16$, respectively) compared to the two scenarios with uncontrollable causes of the misbehavior ($M_{S/UC} = 2.55$, $SD_{S/UC} = 1.23$, and $M_{US/UC} = 2.61$, $SD_{US/UC} = 1.33$, respectively). To test the influence of both stability and controllability statistically, we used separate repeated measures ANOVAs predicting participants' ratings by means

of the manipulated factors (stable vs. unstable; controllable vs. uncontrollable) as within-subjects factors. In line with the descriptive pattern, the stable scenarios received significantly higher stability ratings than the unstable scenarios, $F(1, 232) = 483.57$, $p < .001$, $\hat{\eta}^2_G = .45$. Likewise, the controllable scenarios received significantly higher controllability ratings than the uncontrollable scenarios, $F(1, 232) = 178.39$, $p < .001$, $\hat{\eta}^2_G = .14$.

Direct measure: Endorsement of punishment goals

Overall, special prevention was the most endorsed punishment goal ($M = 3.75$, $SD = 1.33$), whereas retribution ($M = 3.38$, $SD = 1.39$) and general prevention ($M = 3.40$, $SD = 1.47$) were rated as slightly less and equally important. Most important for our hypotheses is the effect of controllability on the endorsement of retribution. However, as is apparent in the left panel of Figure 4, the descriptive picture showed that retribution was rated as roughly equally important across scenarios, with slightly higher ratings in the controllable scenarios (S/C and US/C) than in the uncontrollable scenarios (S/UC and US/UC). To statistically test this pattern, we used a repeated measures ANOVA predicting the endorsement of retribution by the attribution to stable (vs. unstable) and controllable (vs. uncontrollable) causes as within-subjects factors. This revealed significant main effects of both stability, $F(1, 232) = 15.69$, $p < .001$, $\hat{\eta}^2_G = .01$, and controllability, $F(1, 232) = 43.11$, $p < .001$, $\hat{\eta}^2_G = .04$, although effect sizes were actually relatively small.

For the endorsement of utilitarian punishment (i.e., special prevention and general prevention), in turn, we further hypothesized an effect of stability. In contrast to this prediction, however, the endorsement of both special prevention and general prevention was strongly affected by controllability, with higher ratings in the controllable scenarios (S/C and US/C) as compared to the uncontrollable scenarios (S/UC and US/UC). As such, a repeated measures ANOVA on the endorsement of special prevention revealed significant main effects of both

stability, $F(1, 232) = 11.70$, $p = .001$, $\hat{\eta}^2_G = .01$, and controllability, $F(1, 232) = 252.82$, $p < .001$, $\hat{\eta}^2_G = .21$, with an unexpected medium-sized to large effect of controllability. Likewise, there was an unexpected significant medium-sized main effect of controllability on the endorsement of general prevention, $F(1, 232) = 147.06$, $p < .001$, $\hat{\eta}^2_G = .11$, whereas the predicted effect of stability was not significant, $p = .729$, $\hat{\eta}^2_G = .001$.

Indirect measure: Appropriateness of punishment reactions

In contrast to the direct measure of punishment goals, the indirect measure showed that the retributive reaction was rated as most appropriate overall ($M = 2.97$, $SD = 1.62$), followed by the general preventive reaction ($M = 2.36$, $SD = 1.70$), and the special preventive reaction ($M = 2.02$, $SD = 1.73$). Again, most important for our hypotheses is the appropriateness rating of the punishment reactions in the different scenario conditions, which is displayed in the left panel of Figure 5. We once more predicted an effect of controllability on the appropriateness of the retributive reaction and an effect of stability on the appropriateness of the special preventive and general preventive reactions, and analyzed these effects using separate repeated measures ANOVAs for each of the three punishment reactions with the attribution to stable (vs. unstable) and controllable (vs. uncontrollable) causes as within-subjects factors. Mirroring the results for the direct measure of punishment goals, and again in contrast to our hypotheses, ratings of the retributive reaction were comparable across scenario conditions, with slightly higher ratings in the controllable scenarios (S/C and US/C) than in the uncontrollable scenarios (S/UC and US/UC). Again, the ANOVA revealed a significant main effect of controllability on the appropriateness of the retributive punishment reaction, $F(1, 232) = 7.16$, $p = .008$, $\hat{\eta}^2_G = .01$. Although this is in principle in line with the hypothesis, the effect was almost negligible in size. Furthermore, there was no effect of stability on the appropriateness of retributive punishment, $p = .144$, $\hat{\eta}^2_G = .001$.

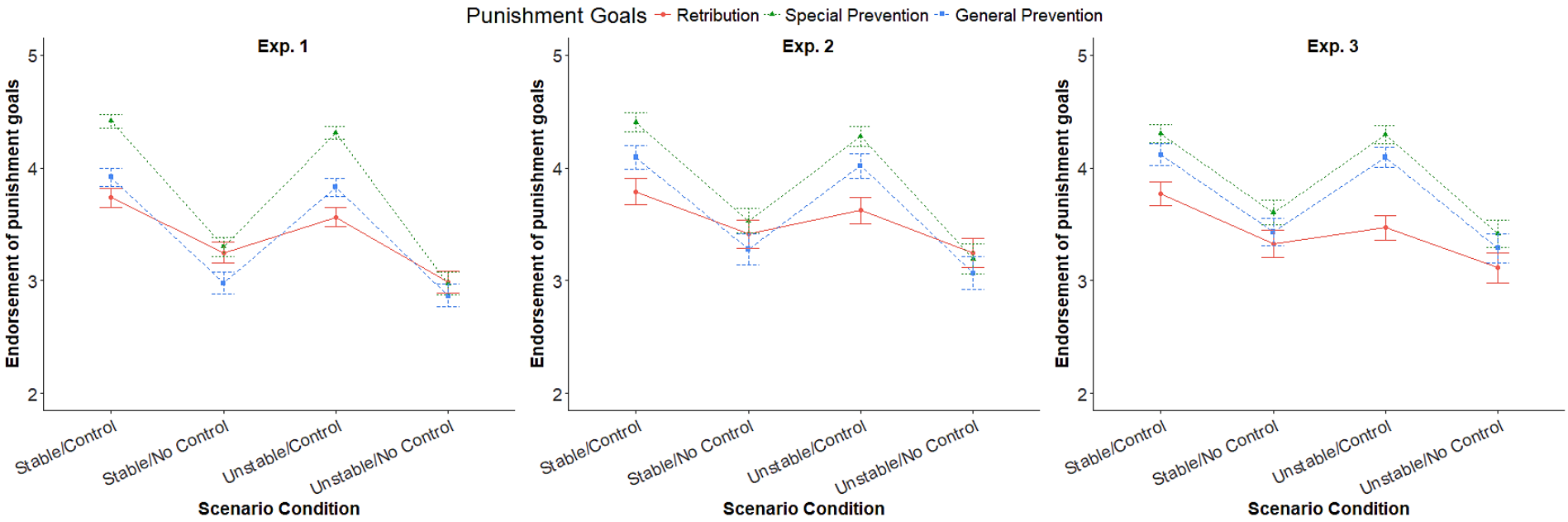

Figure 4. Results of the direct rating of punishment goals in the three experiments: mean endorsement of punishment goals depending on the attribution of the misbehavior in terms of the stability and controllability of its cause. Error bars represent one standard error of the mean.

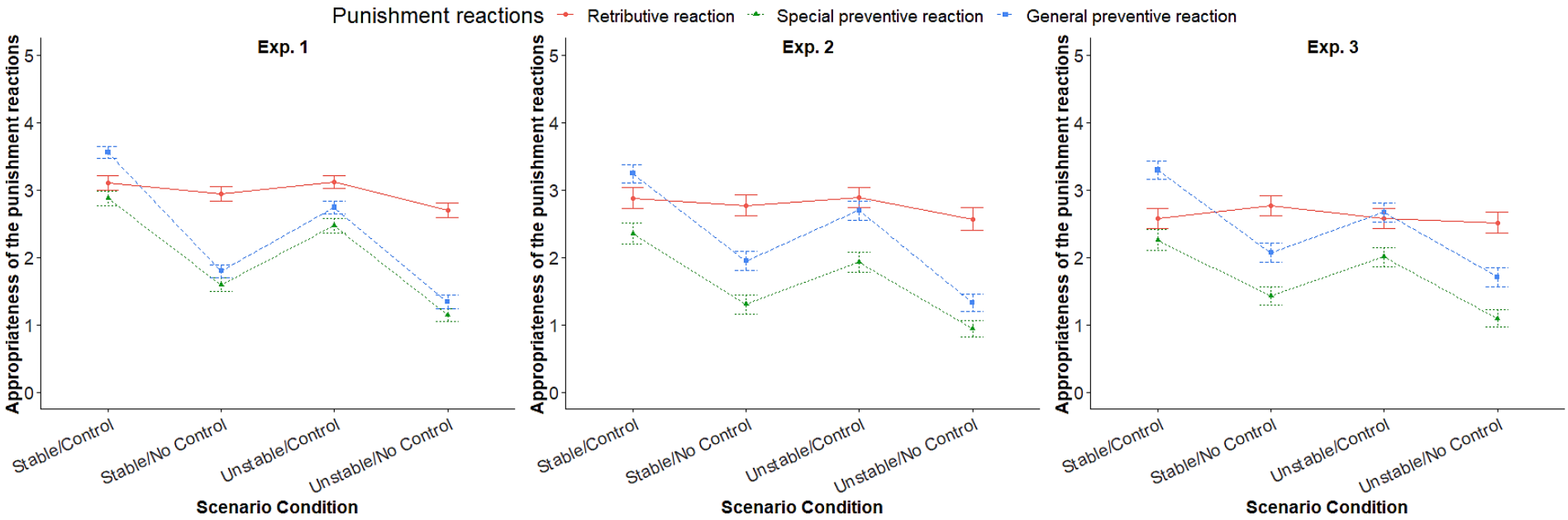

Figure 5. Results of the indirect measurement of punishment goals in the three experiments: mean appropriateness ratings of the three punishment reactions depending on the attribution of the misbehavior in terms of the stability and controllability of its cause. Error bars represent one standard error of the mean.

Also in line with the direct measure, and contrary to the hypotheses, the two utilitarian punishment reactions (i.e., special and general prevention) received higher ratings in the controllable scenarios (S/C and US/C) than in the uncontrollable scenarios (S/UC and US/UC), with substantially larger differences between scenarios. For the appropriateness of the special preventive reaction, the corresponding ANOVA revealed significant main effects of both stability, $F(1, 232) = 27.79$, $p < .001$, $\hat{\eta}^2_G = .02$, and controllability, $F(1, 232) = 224.46$, $p < .001$, $\hat{\eta}^2_G = .15$. Likewise, we found significant main effects of both stability, $F(1, 232) = 62.18$, $p < .001$, $\hat{\eta}^2_G = .04$, and controllability, $F(1, 232) = 321.97$, $p < .001$, $\hat{\eta}^2_G = .22$, on the appropriateness rating of the general preventive reaction. Thus, controllability had the stronger impact on the appropriateness ratings of both utilitarian punishment reactions.

Overall, the results of Experiment 1 were only partially in line with the hypotheses. More specifically, although the perceived controllability of the cause of the student's misbehavior had a significant effect on the support of retributive punishment—with the attribution to a controllable (vs. uncontrollable) cause of the misbehavior leading to a greater support of retribution—effect sizes were small for the direct measure of punishment goals and almost negligible for the indirect measure. Moreover, the attribution to a controllable (vs. uncontrollable) cause of the misbehavior had an unexpectedly strong effect on both the endorsement of special and general prevention as punishment goals (i.e., the direct measure), as well as on the appropriateness ratings of the special and general preventive punishment reactions (i.e., the indirect measure). By contrast, the expected effect of the perceived stability of the cause of the misbehavior was rather small.

4.4 EXPERIMENT 2

In Experiment 1, we found that the attribution of the misbehavior to both stable (vs. unstable) and controllable (vs. uncontrollable) causes had no (or rather negligible) effects on

the approval of retributive punishment in the context of classroom interventions. This was surprising, as past research reported substantial effects of controllability on the support of retribution (Graham et al., 1997; Weiner et al., 1997). Furthermore, it was also unexpected to find that controllability had an effect on both special and general preventive punishment, with the attribution to a controllable (vs. uncontrollable) cause of the misbehavior leading to greater support of special and general prevention. Given these unexpected findings, we considered it vital to test the robustness of the results in a second experiment. Furthermore, results from Experiment 1 came from a convenience sample without any pedagogical background. In turn, given that teaching experience has been found to have an influence on both the attribution of and the interventions following student misbehavior (e.g., Andreou & Rapti, 2010), we recruited a sample of pre-service teachers in Experiment 2.

4.4.1 Methods

Measures and procedure

The experiment was again conducted via the Internet. All manipulations and variables were identical to those used in Experiment 1. Participants were asked to imagine that the students presented in the scenarios were their own students.

Sample

We conducted an a priori power analysis using G*Power (Faul et al., 2009, 2007) to replicate the results found in Experiment 1. Correspondingly, we relied on the smallest effect size observed in Experiment 1 that we considered meaningful, that is, the effect of controllability on the endorsement of retribution as punishment goal ($\hat{\eta}^2_G$ = .04). This analysis revealed the following results: to detect an effect of f = .20 (small to medium-sized) in a repeated measures ANOVA with a power of 1-β = .90 (given α = .05, number of groups = 1, number of measurements = 4, and nonsphericity correction ε = 1), N = 121 participants are necessary.

As mentioned above, we recruited pre-service teachers as participants in Experiment 2. They were approached through mailing lists, social media platforms, and personal contacts. As in Experiment 1, participants had the chance to win one of twenty 20€ gift vouchers as an incentive for participation. Overall, 160 participants started the experiment, of which $N = 119$ (74%) completed it. In this final sample, ages ranged between 18 and 36 years ($M = 23.38$, $SD = 3.09$) and 84% ($n = 100$) of participants were female. These pre-service teachers were in their sixth semester of studies on average ($M = 5.61$, $SD = 3.23$) and mostly studied teaching on high school level (35%), teaching for primary schools (28%), or special education (22%).

4.4.2 Results and discussion

Manipulation check

The manipulation of both the stability and controllability of the cause of the student's misbehavior was again successful. The two scenarios representing stable causes received higher ratings of stability ($M_{S/C} = 3.86$, $SD_{S/C} = 1.13$, and $M_{S/UC} = 3.45$, $SD_{S/UC} = 1.24$, respectively) compared to the scenarios with unstable causes of the misbehavior ($M_{US/C} = 1.67$, $SD_{US/C} = 1.39$, and $M_{US/UC} = 1.26$, $SD_{US/UC} = 1.22$, respectively). Likewise, the two scenarios with controllable causes of the misbehavior received higher ratings of controllability ($M_{S/C} = 3.52$, $SD_{S/C} = 1.23$, and $M_{US/C} = 3.57$, $SD_{US/C} = 1.15$, respectively) compared to the two scenarios with uncontrollable causes of the misbehavior ($M_{S/UC} = 2.76$, $SD_{S/UC} = 1.31$, and $M_{US/UC} = 2.59$, $SD_{US/UC} = 1.38$, respectively). To test the influence of both stability and controllability statistically, we used separate repeated measures ANOVAs predicting participants' ratings by means of the manipulated factors (stable vs. unstable; controllable vs. uncontrollable) as within-subjects factors. In line with the descriptive pattern, the stable scenarios received significantly higher stability ratings than the unstable scenarios, $F(1, 118) = 267.66$, $p < .001$, $\hat{\eta}^2_G = .44$, and the controllable scenarios received significantly higher controllability ratings than the uncontrollable scenarios, $F(1, 118) = 76.28$, $p < .001$, $\hat{\eta}^2_G = .11$.

Direct measure: Endorsement of punishment goals

In line with the results of Experiment 1, special prevention was the most endorsed punishment goal overall (M = 3.85, SD = 1.24), followed by general prevention (M = 3.61, SD = 1.45); retribution (M = 3.52, SD = 1.35) was least endorsed. However, differences between the three goals were smaller as compared to Experiment 1. For the endorsement of the different punishment goals in the scenario conditions, in turn, results were highly similar to Experiment 1, as shown in the middle panel of Figure 4. That is, retribution was again rated as comparably important across scenario conditions, with slightly higher ratings in the controllable scenarios (S/C and US/C) than in the uncontrollable scenarios (S/UC and US/UC). Correspondingly, a repeated measures ANOVA on the endorsement of retribution revealed significant, albeit very small, main effects of both stability, $F(1, 118) = 6.82$, $p = .010$, $\hat{\eta}^2_G = .003$, and controllability, $F(1, 118) = 14.09$, $p < .001$, $\hat{\eta}^2_G = .02$. This is in line with the results of Experiment 1, which also revealed practically negligible effects of the attribution of misbehavior on the endorsement of retribution.

In turn, the two utilitarian punishment goals (i.e., special prevention and general prevention) received higher ratings in the controllable scenarios (S/C and US/C) than in the uncontrollable scenarios (S/UC and US/UC), with considerably larger differences between scenario conditions. Again similar to Experiment 1, the corresponding repeated measures ANOVA showed significant main effects of both stability, $F(1, 118) = 9.21$, $p = .003$, $\hat{\eta}^2_G = .01$, and controllability, $F(1, 118) = 127.33$, $p < .001$, $\hat{\eta}^2_G = .16$, on the endorsement of special prevention, with a stronger impact of controllability. Likewise, the analysis on the endorsement of general prevention once more revealed a significant effect of controllability, $F(1, 118) = 73.94$, $p < .001$, $\hat{\eta}^2_G = .09$, but no effect of stability, $p = .610$, $\hat{\eta}^2_G = .003$.

Indirect measure: Appropriateness of punishment reactions

For the indirect (appropriateness) measure of punishment goals, the pattern was also largely in line with Experiment 1: the retributive reaction was rated as most appropriate overall (M = 2.78, SD = 1.68), followed by the general preventive reaction (M = 2.30, SD = 1.67), and the special preventive reaction (M = 1.63, SD = 1.64). As such, the findings were once again in contrast to those obtained for the direct measure of punishment goals. The middle panel of Figure 5 provides a summary of the appropriateness ratings of the three punishment reactions by the pre-service teachers. As is apparent, the retributive reaction was again rated as rather equally important across scenarios, with slightly higher ratings in the controllable scenarios (S/C and US/C) than in the uncontrollable scenarios (S/UC and US/UC). Correspondingly, the repeated measures ANOVA on appropriateness rating of retribution neither revealed a significant effect of stability, p = .300, $\hat{\eta}^2_G$ < .001, nor of controllability, p = .109, $\hat{\eta}^2_G$ = .004. Effect sizes were negligible and thus highly similar across the two experiments, although reaching a conventional level of statistical significance in the former.

Likewise in line with Experiment 1, both the special preventive reaction and the general preventive reaction received higher ratings in the controllable scenarios (S/C and US/C) than in the uncontrollable scenarios (S/UC and US/UC), with much greater differences between conditions. As such, the main effects of both stability, F(1, 118) = 18.98, p < .001, $\hat{\eta}^2_G$ = .02, and controllability, F(1, 118) = 73.34, p < .001, $\hat{\eta}^2_G$ = .10, were again significant in the repeated measures ANOVA on the appropriateness rating of the special preventive reaction. Similarly, we again found significant main effects of both stability, F(1, 118) = 30.57, p < .001, $\hat{\eta}^2_G$ = .04, and controllability, F(1, 118) = 131.71, p < .001, $\hat{\eta}^2_G$ = .17, in the ANOVA on the appropriateness rating of the general preventive reaction. Thus, as in Experiment 1, controllability rather than stability had the stronger impact on the appropriateness ratings of both utilitarian punishment reactions.

Overall, we found the endorsement of retribution and the appropriateness rating of the retributive punishment reaction to be largely unaffected by the manipulation of both stability and controllability of the cause of the misbehavior. For both measures of special and general prevention, controllability had medium-sized effects, with higher endorsement and appropriateness ratings for the scenarios describing a controllable cause of the misbehavior. In sum, Experiment 2 replicated the findings of Experiment 1 and thus, were again at odds with the hypotheses derived from attribution theory.

4.5 EXPERIMENT 3

As previously summarized, the results of Experiment 1 were replicated in Experiment 2 and thus extended to pre-service teachers. However, a still more rigorous test of whether the observed results also apply for everyday school life arguably requires replication with actual in-service teachers who will have more experience in the classroom and thus with corresponding punishment-related situations. In a third experiment, we therefore aimed to (i.) replicate the main findings of Experiments 1 and 2 and (ii.) to do so in a sample of in-service teachers. In line with our prior findings, we expected retribution to be largely unaffected by the manipulation of both stability and controllability of the cause of the misbehavior as well as an effect of both stability and controllability on the support of special and general prevention as punishment goals, with controllability having the stronger impact.

4.5.1 Methods

Measures and procedure

The experimental design including instructions and materials were identical to Experiments 1 and 2. However, in Experiment 3, participation was possible both via the Internet and through paper and pencil.

Sample

A pragmatic approach was utilized to determine sample size prior to data collection, with a threshold of at least 121 participants (i.e., see a priori power analysis in Experiment 2). Data collection had a planned cessation date, but only if the threshold was reached. Participants (teachers) were recruited individually through mailing lists, social media platforms, and personal contacts, as well as on school level through personal visits and presentation made by the first author. In the latter case, questionnaires were handed out to teachers either personally or they were placed in teachers' mailboxes. To allow for anonymous return of the completed questionnaires, a box was provided in the schools. We obtained a 25% response rate, which is comparable to similar field studies (e.g., Reyna & Weiner, 2001).

As an incentive for participation, teachers had a chance to win one of twenty 20€ gift vouchers. Furthermore, each participating school had a chance to win a large-scale voucher for school supplies. By the end of our survey period, 103 teachers started the online-version of the experiment, with 74 (72%) completing it. Additionally, 67 teachers completed the paper and pencil version of the experiment through the recruitment in schools. The total sample therefore comprised N = 141 participants. Around two thirds of participants (i.e., n = 92; 65%) were female (two participants did not indicate their gender), and ages ranged between 23 and 70 years (M = 40.77, SD = 11.13).

According to a post-hoc power analysis conducted with G*Power (Faul et al., 2009, 2007), this results in highly satisfactory power of $1-\beta = .94$ to detect a small to medium-sized effect of $f = .20$ in a repeated measures ANOVA (given $\alpha = .05$, number of groups = 1, number of measurements = 4, and nonsphericity correction $\varepsilon = 1$).

4.5.2 Results and discussion

Manipulation check

Again, the manipulation of both the stability and controllability of the cause of the student's misbehavior was successful. The two scenarios representing stable causes received

higher ratings of stability ($M_{S/C}$ = 3.97, $SD_{S/C}$ = 1.09, and $M_{S/UC}$ = 3.46, $SD_{S/UC}$ = 1.17, respectively) compared to the scenarios with unstable causes of the misbehavior ($M_{US/C}$ = 1.60, $SD_{US/C}$ = 1.29, and $M_{US/UC}$ = 1.10, $SD_{US/UC}$ = 1.11, respectively). Likewise, the two scenarios with controllable causes of the misbehavior received higher ratings of controllability ($M_{S/C}$ = 3.19, $SD_{S/C}$ = 1.33, and $M_{US/C}$ = 3.49, $SD_{US/C}$ = 1.21, respectively) compared to the two scenarios with uncontrollable causes of the misbehavior ($M_{S/UC}$ = 2.60, $SD_{S/UC}$ = 1.23, and $M_{US/UC}$ = 2.92, $SD_{US/UC}$ = 1.47, respectively). However, this effect was slightly smaller than in Experiments 1 and 2. To test the influence of both stability and controllability statistically, we used separate repeated measures ANOVAs predicting participants' ratings by means of the manipulated factors (stable vs. unstable; controllable vs. uncontrollable) as within-subjects factors. In line with the descriptive pattern, the stable scenarios received significantly higher stability ratings than the unstable scenarios, $F(1, 135) = 444.04$, $p < .001$, $\hat{\eta}^2_G = .51$, and the controllable scenarios received significantly higher controllability ratings than the uncontrollable scenarios, $F(1, 136) = 34.83$, $p < .001$, $\hat{\eta}^2_G = .05$.

Direct measure: Endorsement of punishment goals

Overall, results from the direct (endorsement) measure of punishment goals were again in contrast to the initial hypothesis, but largely replicated the findings of Experiments 1 and 2. That is, special prevention was again the most endorsed punishment goal (M = 3.90, SD = 1.23), followed by general prevention (M = 3.73, SD = 1.36), and retribution (M = 3.42, SD = 1.41). As depicted in the right panel of Figure 4, endorsement of retribution was rated as rather equally important across conditions, with highest ratings in the S/C scenario. As in Experiment 1 and 2, the effects of both stability, $F(1, 138) = 15.22$, $p < .001$, $\hat{\eta}^2_G = .01$, and controllability, $F(1, 138) = 24.84$, $p < .001$, $\hat{\eta}^2_G = .02$, were once more small in size, yet statistically significant in a repeated measures ANOVA on the endorsement of retribution as punishment goal.

In turn, differences across scenarios for special prevention and general prevention were more substantial, with higher ratings in the controllable scenarios (S/C and US/C) than in the uncontrollable scenarios (S/UC and US/UC). As such, we again found a significant main effect of controllability on the endorsement of special prevention, $F(1, 138) = 91.23$, $p < .001$, $\hat{\eta}^2_G = .10$. However, the initially hypothesized effect of stability on the endorsement of special prevention, as found in our first two experiments, was not replicated, $p = .286$, $\hat{\eta}^2_G = .002$. Furthermore, analyses again revealed a significant effect of controllability on the endorsement of general prevention, $F(1, 138) = 64.19$, $p < .001$, $\hat{\eta}^2_G = .08$, but no effect of stability, $p = .256$, $\hat{\eta}^2_G = .001$.

Indirect measure: Appropriateness of punishment reactions

Similar to the results of our first two experiments, the retributive reaction was rated as most appropriate overall ($M = 2.61$, $SD = 1.76$), followed by the general preventive reaction ($M = 2.44$, $SD = 1.75$), and the special preventive reaction ($M = 1.70$, $SD = 1.71$). The appropriateness ratings of the three punishment reactions in the four scenario conditions are depicted in the right panel of Figure 5. Replicating Experiments 1 and 2, there were no differences between the scenario versions for the appropriateness rating of the retributive reaction. This is also reflected in the results of the corresponding repeated measures ANOVA, revealing that there was neither a significant effect of stability, $p = .250$, $\hat{\eta}^2_G = .001$, nor of controllability, $p = .588$, $\hat{\eta}^2_G < .001$, on the appropriateness rating of the retributive punishment reaction.

By contrast, but also in line with Experiments 1 and 2, the two utilitarian punishment reactions (i.e., the special preventive reaction and the general preventive reaction) received higher ratings in the controllable scenarios (S/C and US/C) than in the uncontrollable scenarios (S/UC and US/UC). More specifically, we found significant main effects of both stability, $F(1, 138) = 8.54$, $p = .004$, $\hat{\eta}^2_G = .01$, and controllability, $F(1, 138) = 67.55$, $p < .001$, $\hat{\eta}^2_G = .07$, on

the appropriateness rating of the special preventive reaction, although the effect of controllability was slightly smaller in Experiment 3 as compared to the first two experiments. Likewise, a repeated measures ANOVA on the appropriateness rating of the general preventive reaction revealed significant main effects of both stability, $F(1, 138) = 19.74$, $p < .001$, $\hat{\eta}^2_G = .02$, and controllability, $F(1, 138) = 95.89$, $p < .001$, $\hat{\eta}^2_G = .10$, and showed that perceived controllability again had the highest effect on the appropriateness rating of the general preventive reaction.

In sum, Experiment 3 largely replicated the findings from Experiments 1 and 2—with a sample that arguably has more experience in the classroom and thus with corresponding punishment-related situations: in-service teachers. More specifically, we again found that retribution was basically unaffected by the manipulation of both stability and controllability of the cause of the misbehavior. By contrast, both the perceived stability and controllability of the cause of the misbehavior influenced the support of special and general preventive punishment, with controllability having the stronger impact.

4.6 GENERAL DISCUSSION

Past research has shown that the goals individuals intend to achieve when punishing an offender are influenced by the perceived stability and controllability of the cause of the misbehavior (Graham et al., 1997; Weiner et al., 1997). That is, the support of utilitarian punishment goals is subject to the perceived stability of the cause of the misbehavior. Likewise, the support of retribution as a punishment goal is subject to the perceived controllability of the cause of the misbehavior. In the present research, we applied this attributional model of punishment goals to study punishment behavior in an environment that is fundamental to individual development (Wald & Losen, 2003), and in which punishment decisions are realized on a daily basis: the school (Kulinna et al., 2006). More precisely, lay participants (Experiment 1), pre-service teachers (Experiment 2), and in-service teachers (Experiment 3) read several versions of a scenario describing a student destroying the belongings of another student. Using a 2 x 2 within-

subjects design, we manipulated the stability (stable vs. unstable) and controllability (controllable vs. uncontrollable) of the cause of the misbehavior described in the scenarios. We tested whether this manipulation has an influence on the support of three punishment goals, that is, retribution, special prevention, and general prevention. These punishment goals were measured both directly (i.e., asking for the endorsement of punishment goals) and indirectly (i.e., asking for the appropriateness of concrete, pretested punishment reactions), as past research has shown the potential shortcomings of an exclusive direct measure of punishment goals (e.g., Carlsmith, 2008; Crockett et al., 2014; Goodwin & Benforado, 2015).

Across all three experiments, we consistently found that the manipulation of attribution had a strong influence on the endorsement of utilitarian punishment (i.e., both special and general prevention). However, although past research has shown that stability is the key factor for the support of utilitarian punishment (e.g., Graham et al., 1997; Weiner et al., 1997), it was the manipulation of the controllability of the cause of the misbehavior that was crucial in our data referring to a classroom setting. More precisely, utilitarian punishment goals—and reactions achieving these—received more support when the cause of the misbehavior was perceived as controllable rather than uncontrollable. This is in line with research in education showing that controllability is a key factor for punishment in general, with generally milder penalties whenever the cause of the misbehavior is perceived as uncontrollable (Soodak & Podell, 1994; Tollefson, 2000). In turn, the unexpectedly negligible effect of stability on the support of utilitarian punishment may be interpreted as a hopeful sign that teachers generally assume that all students can change, even though their misbehavior is subject to a rather stable cause (as ensured by the manipulation checks).

Despite the effect of the perceived cause of the misbehavior on the support of utilitarian punishment, there was a rather negligible effect of the manipulation of attribution on both the endorsement of retribution and the appropriateness rating of the retributive reaction. This is

surprising on first sight, given that past research has found an effect of the perceived controllability of the misbehavior on the support of retribution (e.g., Graham et al., 1997; Weiner et al., 1997). However, these surprising results may be explained by differences in the conceptualization of retribution: whereas past research applied a rather unidimensional conception of retribution—focusing on "retribution as revenge"—retribution as "just deserts" also implies a focus on the restoration of justice (Gerber & Jackson, 2013). In other words, whereas giving the student a negative feedback because "it is what she deserved" may be a working conceptualization of retribution in the case of an exam failure, this definition does not extend to the case of punishing student misbehavior once a victim is involved. In this latter case, retribution aims at both restoring justice and giving the offender what she deserves for her misbehavior. This more comprehensive definition of retributive punishment as a way to restore justice may be the reason why the support for this punishment goal was mostly independent of the cause of the misbehavior in our experiments.

Although the general pattern of the association between attribution of the cause of misbehavior and participants' punishment goals were mostly comparable across the direct and indirect measure of punishment goals, there were also some consistent differences: participants indicated a greater endorsement of utilitarian punishment (in particular of special prevention) in the direct measure of punishment goals. This appears to contrast with a vast amount of research suggesting that retribution is generally the primary goal of punishment (e.g., Carlsmith et al., 2002). However, our finding is in line with more recent results indicating the influence of status and power on laypeople's punishment goals, with more powerful people having a greater focus on utilitarian punishment goals (Mooijman et al., 2015). Given that teachers may also be considered as powerful, this may explain their preference for utilitarian punishment. However, and in contrast to the direct endorsement rating of punishment goals, the more indirect (appropriateness) measure of punishment goals revealed higher scores for retributive pun-

ishment than for utilitarian punishment. This divergence is once more consistent with the hypothesis that individuals are arguably not aware of the goals that actually drive their punishment behavior (Applegate, Cullen, Turner, & Sundt, 1996) and, in particular, that they overestimate the influence of utilitarian goals on their punishment decisions (Carlsmith, 2008; Crockett et al., 2014).

Furthermore, we introduced general prevention as an important goal of punishment to a literature that has almost exclusively focused on the dichotomy between retribution and utilitarianism (e.g., Darley, Carlsmith, & Robinson, 2000)—with the latter most often being restricted to the prevention of recidivism or, as herein referred to, special prevention (although occasionally, the primary subject of utilitarian punishment is the prevention of imitation; e.g., Mooijman et al., 2017). Although causal attribution had similar effects on special prevention and general prevention, our data suggests that general prevention is of great importance for punishment behavior, at least in the context of classroom interventions. In fact, in all experiments, the general preventive reaction was rated as most appropriate whenever misbehavior was attributed to a stable and controllable cause. In any case, our results imply that these two utilitarian punishment goals differ to some extent, and future research is needed to illuminate this distinction further.

From a more applied perspective, the present research is the first to analyze teachers' punishment goals from a social psychological perspective. Whereas justice has already been established as a major topic in educational research (e.g., Donat et al., 2012; Peter & Dalbert, 2010), there has been, to the best of our knowledge, only a single attempt to analyze teachers' behavior applying an attributional model of decision making (Reyna & Weiner, 2001). However, that study concentrated on teachers' goals when giving feedback following a failed student exam rather than on punishment in classroom settings. By contrast, as represented in the present research, it is not the evaluation of student performance, but the treatment of student misbehavior which is a main stressor for both teachers (Melnick & Meister, 2008) and students (Fan &

Chan, 1999; Israelashvili, 1997). Our research provides first evidence for the primary goals of teachers' punishment behavior as a function of the attribution of student misbehavior.

This first insight into the motivational basis of teachers' punishment may serve as a starting point for more research on the (fair) treatment of student misbehavior. Future research along these lines should not only deal with the teachers' perspective on classroom interventions, but also consider the students' view on teachers' punishment behavior (Lewis, 2001). In case of teachers' punishment goals, students may arguable perceive a punishment reaction aiming to achieve general prevention as less fair or legitimate, which could even lead to a decrease in students' rule compliance (Mooijman et al., 2017). Moreover, the results can enhance our knowledge about teachers' view on classroom intervention strategies and, as such, be incorporated into pre-service teacher education and in-service teacher professional development programs to decrease teachers' insecurity in dealing with student misbehavior (Melnick & Meister, 2008). Therefore, one first step may be to increase teachers' awareness of the attributional processes taking place when perceiving student misbehavior. This may help them to reflect on their own behavior and the goals they attend to achieve by punishment.

Notwithstanding these advantages, some potential limitations of the present research must be acknowledged. First, there is arguably an inherent difference between imagining a situation and actually experiencing it (Hughes & Huby, 2004; Schoenberg & Ravdal, 2000). In the present experiments, participants were asked to imagine a specific misbehavior using a scenario. The quality of such scenarios—that is, whether participants are able to feel involved in the described situation—largely depends on whether participants experience it as relevant and authentic (e.g., Hughes, 1998). We therefore put great effort into the development of our material, consulting teachers as experts to evaluate and improve the scenario and the potential reactions of the teachers. Furthermore, such an experimental approach has already found to be successful in examining teachers' evaluation and behavior in other domains (e.g., Baudson & Preckel, 2013). Nonetheless, future research may consider studying actual student misbehavior

and punishment behavior by teachers in school settings using observations in the field (Klein, 2008; Klein, Orasanu, Calderwood, & Zsambok, 1993; Lipshitz, Klein, Orasanu, & Salas, 2001).

Another potential limitation associated with the scenario used is that we confronted participants with a single scenario describing one specific case of student misbehavior (i.e., a student destroying the belongings of another student). Thus, the results may be subject to the specificity of this misbehavior and the reactions applied in our experiments. For instance, people's punishment goals may also be influenced by other aspects of the misbehavior as studied herein, such as the magnitude of harm caused (Carlsmith, 2006). Future research may therefore conceptually replicate and extend the present findings to more diverse forms of student misbehavior.

In conclusion, we found effects of the attribution to a controllable (vs. uncontrollable) cause of a student's misbehavior on the support of special prevention and general prevention as punishment goals. By contrast, the teachers' endorsement of retribution appeared to be independent of the perceived cause of the misbehavior. Notwithstanding the implications of these results, future research is needed to investigate potential consequences of such goals. For example, given that other individuals are typically involved in situations of misbehavior (Schmitt et al., 2005), it is plausible to assume that these others (e.g., a victim) want to see the offender receiving her just deserts (i.e., retribution), independent of the controllability of the cause of the misbehavior. Future research may investigate how victims interpret a teacher's punishment behavior if this is primarily geared to achieve utilitarian goals. Possibly, such discrepancies may trigger feelings of injustice for the victim, even though punishment has taken place. We hope that our findings encourage research in school settings along these lines.

5. PROJECT 3: "APPROPRIATE BUT UNFAIR? COMPARING TEACHERS' AND STUDENTS' PERSPECTIVES ON THE TREATMENT OF STUDENT MISBEHAVIOR"

5.1 ABSTRACT

The treatment of student misbehavior is both a major challenge for teachers and a potential source of students' experiences of injustice in school. By implication, it is vital to understand teachers' treatment of student misbehavior vis-à-vis students' perceptions. One key dimension of punishment behavior reflects the underlying motivation and goals of the punishment. In the present research, we investigated the perspectives of both teachers and students concerning the purposes of punishment. Specifically, we were interested in the extent to which teachers and students show preferences for either retribution (i.e., evening out the harm caused), special prevention (i.e., preventing recidivism of the offender), or general prevention (i.e., preventing imitation of others) as punishment goals. Therefore, teachers (N = 260) and students around the ages of 10 and 11 (N = 238) were provided with a scenario depicting a specific student misbehavior. Participants were asked to indicate their endorsement of the three goals as well as to rate different punishment reactions known to primarily achieve one specific goal but not the other two. Results show that teachers largely prefer general prevention, whereas students rather support special prevention and retribution. This discrepancy was particularly large in participants' ratings of specific punishment reactions, whereas differences of teachers' and students' general endorsement of punishment goals were comparably small. Overall, the present research contributes to the development of classroom intervention strategies that may reduce conflicts in student-teacher-interactions.

Appropriate but unfair? Comparing teachers' and students' perspectives on the treatment of student misbehavior

In February 2018, the case of a school in Germany went through the media because the teachers of this school wrote an open letter to the parents, disclosing that they considered themselves incapable of further teaching their students due to the extent of student misbehavior ("Grundschule in Sachsen-Anhalt", 2018). More precisely, teachers reported that students were violent against each other, displayed no respect to the teachers, and ignored any rules. Strikingly, these were not teachers from a secondary school in an urban, underprivileged district, but from an elementary school in a small, rural town with students from all social classes. Unfortunately, this incident is by no means an exception, as recent official statistics corroborate that the frequency of student misbehavior in German schools is increasing (e.g., Landeskriminalamt Nordrhein-Westfalen, 2017).

As this example shows, there is a wide range of student misbehaviors that occur in class on a daily basis (e.g., Kulinna et al., 2006; Wheldall & Merrett, 1988). Dealing with such incidents and finding appropriate classroom intervention strategies is a crucial challenge and major concern for both practitioners and scholars. In fact, there is a notable body of evidence suggesting that student misbehavior is a primary cause of stress for teachers and is also largely responsible for the high prevalence of burnout syndrome in teachers (e.g., Aloe et al., 2014; Brouwers & Tomic, 2000). In the best case, student misbehavior can be prevented by positive and proactive classroom management approaches (e.g., Sugai & Horner, 2006). However, sometimes such misbehavior cannot be prevented by a well-prepared lesson (e.g., if it occurs during break). Nonetheless, teachers are typically forced to deal with such incidents—and it is a major concern for teachers to find suitable and effective interventions in such situations (Melnick & Meister, 2008).

Likewise, students' perceptions of injustice in school is of great importance, given its noteworthy influence on class climate (Peter & Dalbert, 2010), the extent of bullying behavior (Donat et al., 2012), as well as the students' well-being (Kamble & Dalbert, 2012), their academic self-concept, motivation, and grades (Peter et al., 2012). Importantly, students' experiences of injustice in school are not limited to and indeed is only marginally driven by the grading or evaluation of students' performances. Instead, the treatment of student misbehavior appears to be an important factor (Israelashvili, 1997). In fact, research suggests that situations of punishment are among the most frequently reported situations of students' experiences of injustice, even excluding situations of false allegations (Fan & Chan, 1999). In other words, even if the incident of student misbehavior and the offender appear to be clear, students frequently feel mistreated by their teachers.

Importantly, the "wrong" treatment of student misbehavior may have a fundamental impact on students' lives, with the "school to prison pipeline"—characterizing the suggestion that individual school suspension rates are correlated with the probability to be sent to jail as an adult (Wald & Losen, 2003)—as the most extreme case. Therefore, it is vital not only to study and understand teachers' punishment behavior, but also to consider the students' perspective on teachers' reactions to student misbehavior. As a result, one may understand what characteristics of teachers' strategies in classroom discipline cause students' experiences of injustice in such situations. In the long run, such knowledge will likely aid the design of school policies and teacher interventions.

One core aspect of teachers' intervention strategies concerns the goals they intend to achieve when reacting to student misbehavior. According to a long tradition of social psychological research, individuals' punishment decisions are generally driven by two major goals: *retribution* and *utilitarianism* (e.g., Carlsmith, 2006; Gromet & Darley, 2009a). In a retributive approach, punishment is backward-oriented and primarily intended to paying back harm doers

for their misconduct (Carlsmith, 2006; Kant, 1952). Accordingly, the primary goal of punishment is to rebalance the wrong that has been committed by the offense and, thus, to restore justice (Gerber & Jackson, 2013). By contrast, in a utilitarian approach, punishment is future-oriented and primarily intended to prevent the offender and potential would-be offenders from committing future norm violations (Bentham, 1962; McCullough et al., 2013). As such, utilitarian punishment can further be differentiated into *special prevention* (i.e., preventing further misconduct by the perpetrator herself) and *general prevention* (i.e., preventing future misconduct by unrelated others; Keller et al., 2010; Rucker et al., 2004).

Most research investigating people's punishment behavior suggests that it is predominantly driven by retribution (i.e., to even out the wrong that has been done) rather than special or general prevention (i.e., to prevent future crimes by the offender or others; Carlsmith, 2006; Carlsmith et al., 2002), leading to an apparent consensus that "people are intuitive retributivists" (Carlsmith & Darley, 2008, p. 211). However, this very general result cannot easily be translated to teachers' punishment behavior in school, given that leaders (i.e., people in a powerful position) show greater preferences for more utilitarian punishment goals, whereas subordinates show greater preferences for more retributive punishment goals (Mooijman et al., 2015).[6] Therefore, it appears plausible that there is an inherent difference between the teachers'—as "leaders" in the school context—punishment goals (and their reaction achieving these goals) and the students'—as "subordinates" in the school context—perception of what may be a reasonable goal and reaction to be shown in a given situation. Correspondingly, this difference—if indeed present—may intensify students' experiences of injustice, even if punishment of the misbehavior is generally justified.

[6] In this research, utilitarian punishment was mainly defined in terms of preventing future misbehavior by observers and, thus, largely represented general prevention.

5.2 THE PRESENT RESEARCH

Given that the treatment of misbehavior is both a major challenge for teachers (Melnick & Meister, 2008) and a potential source of negative perceptions and reactions by students, especially if they feel mistreated (Fan & Chan, 1999; Israelashvili, 1997), it is vital to understand teachers' treatment of student misbehavior vis-à-vis students' perceptions. Therefore, the goals of the present research are twofold.

First, we investigated which punishment goals teachers generally endorse and how they rate specific punishment reactions that serve ultimately different goals. Past research has shown that being in a more powerful position is positively associated with the support of general prevention (Mooijman et al., 2015). Teachers may be considered as "leaders" in the school context (Goddard, 2000) and, thus, we hypothesized that teachers show greater preferences for general prevention compared to retribution. Given the natural overlap of special prevention and general prevention as two related, yet different dimensions of utilitarian punishment, we also expected that teachers generally support special prevention over retribution, whereas we had no hypothesis regarding potential differences between special and general prevention. It should be noted that any hypotheses concerning the role of special prevention were comparatively more speculative, as past research did not explicitly test the influence of power and hierarchy on the support of special prevention.

Second, we investigated the extent to which students share the teachers' preferences in punishment goals and their perception of the punishment reactions pursuing specific goals. Students may be considered as "subordinates" in the school context and, thus, should show less preference for utilitarianism (Mooijman et al., 2015) but rather a general preference for retribution (e.g., Carlsmith et al., 2002; Gromet & Darley, 2009a). We therefore hypothesized that students should indicate a higher endorsement of retribution and should prefer a teacher's reaction to student misbehavior that is designed to achieve a more retributive punishment goal. Most

importantly, the above hypotheses regarding teachers' and students' support of the three punishment goals imply the existence of differences in the punishment goal preferences of the two groups, both in their general endorsement of the goals and in their judgment of reactions pursuing these goals. In other words, our core prediction involves a pattern that yield clear differences in how teachers vs. students evaluate punishment goals and corresponding reactions to misbehavior.

To test these hypotheses, we provided (pre-service and in-service) teachers and students around the ages of 10 and 11 with a scenario describing a specific student misbehavior. We examined their preferences for retribution, special prevention, and general prevention as the primary goal of punishing the misbehaving student, that is, whether either of the goals is more supported than the other goals and whether these preferences are compatible across teachers and students. Importantly, past research has revealed substantial differences between the general endorsement of the punishment goals and actual punishment behavior achieving these goals (e.g., Carlsmith, 2008; Crockett et al., 2014). Therefore, we first investigated and compared teachers' and students' preferences regarding their general endorsement of the punishment goals in the given scenario. Second, we asked teachers to rate the appropriateness of three punishment reactions that were shown (from pretests) to primarily serve one of the goals but not the other two. We then compared teachers' preferences in this reaction appropriateness measure to students' rating of these punishment reactions regarding how fair, appropriate, and just they were, if shown by a teacher.

As recommended, we report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures (Simmons et al., 2012). Furthermore, all materials (including the instructions and materials used in our preliminary study, along with all data, analyses scripts and supplementary analyses) are available at the OSF and can be accessed during double-blind peer-review via the following link:

https://osf.io/r5d8v/?view_only=9fe8285bbc7f435b902eb88b6b8571dd.

5.2.1 Methods

Pretest of the material: Test of reactions

Past research has revealed rather weak correlations between participants' explicit endorsement of punishment goals and their actual punishment behavior (e.g., Crockett et al., 2014). Specifically, it appears that people indicate a rather high endorsement of all punishment goals and generally overestimate the influence of utilitarian goals on their own punishment decisions, whereas their actual punishment behavior is mostly retributive (Carlsmith, 2008; Crockett et al., 2014). Thus, a "direct" endorsement measure of punishment goals may be biased in favor of utilitarian punishment. Consequently, it is important to consider additional, less direct measures of punishment goals. A more indirect measure of punishment goals provides participants with specific punishment reactions that serve certain goals without stating these goals explicitly and asks them to indicate their judgment of these reactions. To the extent that the reactions are known to differ on how well they serve different punishment goals (e.g., from pretests), one can consider participants' ratings as an indirect measure of punishment goals that does not require individuals to introspectively predict the exact goals they would intend to achieve.

Therefore, we created several possible teacher reactions to the student misbehavior. These reactions were pretested for the degree to which they served each of the punishment goals: retribution, special prevention, and general prevention. We aimed to identify three reactions (one for each punishment goal) that would primarily serve one of the goals but not the other two. Additionally, the reactions were intended to be perceived as equally severe.

To pretest the reactions, we gave $N = 122$ participants definitions of the three punishment goals and asked them to rate the extent of goal achievement for the nine reactions to the student misbehavior. Thus, we were able to extract three punishment reactions that were perceived as primarily serving one of the goals. These reactions can then be used to measure teachers' support of the punishment goals more indirectly, as well as students' judgment of potential

teacher reactions that are known to differ in the goals they intend to achieve. Material, data and detailed results (including test statistics) of this preliminary study can be found in detail in the supplementary material at the OSF.

Measures and procedure

For pre-service and in-service teachers, data collection took place both via the Internet and through paper and pencil. After providing informed consent, participants received a scenario describing a student destroying a recently prepared hand drum of another student. This scenario read as follows:

> "Within the past lessons, you manufactured hand drums with your students that you plan to use today. You briefly turn to the board. Once you face the class again, you see how **Florin** causes a hole in **Maxi's drum**. As a result, the drum is broken."[7]

Participants were then asked to answer several control variables regarding the perception of the misbehavior, starting with (i.) the *stability* and (ii.) *controllability* of the cause of the student's misbehavior. Furthermore, participants were asked to indicate (iii.) the student's *responsibility* for what occurred, how much (iv.) *anger* and (v.) *sympathy* they would feel towards the misbehaving student as well as (vi.) to what degree it is possible to *influence* the student's future behavior. Each response was rated on a 6-point scale ranging from 0 = "*not at all*" to 5 = "*completely*." Next, they answered the more "indirect" measure of punishment goals and rated the *appropriateness* of each of the three punishment reactions extracted from our pre-study (see above), each reflecting a different punishment goal. These reactions are provided in Table 1 (see p. 56). Again, each reaction was rated using a 6-point scale ranging from 0 = "*not at all appropriate*" to 5 = "*completely appropriate*." Finally, participants provided their general *endorsement* of the three punishment goals in the specified situation (i.e., the direct measure of

[7] The names of the students in the scenario were chosen to be gender neutral.

punishment goals), indicating the goals they would want to accomplish if presented a chance or obliged to react to the misbehaving student. To this end, we adapted one item for each punishment goal (retribution, special prevention, and general prevention) from Orth (2003) and Weiner and colleagues (1997). The item measuring endorsement of retribution read as follows: "To what extent would you like to react to even out the wrong that Florin has done?"; the item measuring endorsement of special prevention read "To what extent would you like to react to prevent recidivism by Florin?"; and the item measuring endorsement of general prevention read "To what extent would you like to react to prevent other students of showing similar behavior in the future?" Each item was answered on a 6-point scale ranging from 0 = "*not at all*" to 5 = "*completely*," with higher values indicating stronger endorsement of a particular punishment goal. After answering all questions, participants worked on several other scenarios that pertained to a different, unrelated research question. Thus, this part of the material will not be further discussed in the present article. Finally, participants provided demographic information, before they were fully debriefed and thanked.

For students, data collection took place in the fifth and sixth grade of three German public schools. On arrival in the classroom, students provided the experimenters with the consent form signed by their parents or legal guardians. Then, students were seated individually in front of computers. Two experimenters welcomed the class and gave them verbal instructions on the subsequent tasks. The first task was on an unrelated research question and will therefore not be further discussed in the present article. Subsequently, students provided demographic information before they were introduced to the main task. Therein, students were asked to indicate their perceptions of a misbehavior and different reactions to it. This misbehavior and the reactions were presented in short comic strips. Using an exemplary misbehavior, students received comprehensive instructions on the procedure of the task. They were then introduced to the main protagonists of the comics: two students (a girl and a boy) and a teacher (either female

or male). After receiving these instructions, students were asked to work on the task individually. They were provided with a comic strip depicting one student destroying the recently prepared hand drum of the other student (i.e., the same scenario that teachers received). The gender of the misbehaving student and the victim was counterbalanced (i.e., it was either the boy destroying the drum of the girl or vice versa). Next, students received three comic strips depicting potential reactions of the teacher in the comic (again, the teacher was either female or male, counterbalanced). They were asked to rate the extent to which they perceived these reactions as *just*, *appropriate*, and *fair* on 6-point scales ranging from 0 = "*not at all*" to 5 = "*completely*." Subsequently, they answered all of the control questions outlined above (i.e., the *stability* and *controllability* of the cause of the student's misbehavior, the student's *responsibility* for what occurred, how much *anger* and *sympathy* they would feel towards the misbehaving student, and to what degree it is possible to *influence* the student's future behavior), as well as their general *endorsement* of the three punishment goals in the specified situation (i.e., the direct measure of punishment goals). Finally, students were fully debriefed and thanked.

Sample

Data from both (pre-service and in-service) teachers and students were collected in conjunction with projects on other research questions. Thus, we conducted an a priori power analysis for the planned comparisons of teachers' and students' punishment goal preferences using G*Power (Faul et al., 2009, 2007), but only to determine a lower boundary for the data collection (i.e., a threshold to stop data collection). The power analysis was specified to detect a within-between interaction in a mixed model with two groups (teachers and students) as between-subjects factor and the three punishment goals (retribution, special prevention, and general prevention) as within-subjects factor. To detect a small to medium-sized effect of $f = .15$ with a high power of $1-\beta = .90$ (given $\alpha = .05$, number of groups = 2, number of measurements

= 3, and nonsphericity correction ε = 1), *N* = 286 participants (i.e., 143 teachers and 143 students) are required.

For pre-service and in-service teachers, data collection had a planned cessation date, but only if the threshold of approximately 145 participating teachers was reached. To enable data collection of students, we collaborated with several public schools and were able to recruit three schools to participate in our project before data collection started. Such collaborations entail that it is not possible to simply stop data collection once reaching the required sample size. Thus, we planned to collect data in three schools and to recruit further schools for participation if we did not reach the threshold of approximately 145 participating children.

Pre-service and in-service teachers were recruited individually through mailing lists, social media platforms, and personal contacts, as well as on school level through personal visits and presentation made by the first author. In the latter case, questionnaires were handed out to teachers either personally or they were placed in teachers' mailboxes. To allow for anonymous return of the completed questionnaires, a box was provided in the schools. We obtained a 25% response rate, which is comparable to similar field studies (e.g., Reyna & Weiner, 2001). As an incentive for participation, both pre-service and in-service teachers had a chance to win one of twenty 20€ gift vouchers. Furthermore, each participating school had a chance to win a large-scale voucher for school supplies.

For pre-service teachers, 160 participants started the experiment, of which *N* = 119 (74%) completed it. In this final sample of pre-service teachers, ages ranged between 18 and 36 years (*M* = 23.38, *SD* = 3.09) and 84% (*n* = 100) of participants were female. These pre-service teachers were in their sixth semester of studies on average (*M* = 5.61, *SD* = 3.23) and mostly studied teaching on high school level (35%), teaching for primary schools (28%), or special education (22%). For in-service teachers, 103 participants started the online-version of the experiment, with *n* = 74 (72%) completing it. Additionally, *n* = 67 teachers completed the paper and pencil version of the experiment through the recruitment in schools. The total sample of in-

service teachers therefore comprised N = 141 participants. Around two thirds of these participants (i.e., n = 92; 65%) were female (two participants did not indicate their gender), and ages ranged between 23 and 70 years (M = 40.77, SD = 11.13). In total, we collected complete data sets from N = 260 pre-service and in-service teachers, of whom n = 192 were female (74%).

For students, we collected data in the fifth and sixth grade of three public German schools. After 12 sessions, N = 238 children had participated in the study. Around 45% of participants (i.e., n = 106) were female, most children's mother language was German (92%), and ages ranged between 9 and 12 (M = 10.46, SD = 0.61; one child did not indicate any age).

5.2.2 Results

Before conducting in-depth analyses on our main variables of interest (i.e., the punishment goal preferences of teachers and students), we tested whether pre-service and in-service teachers differed at all in their support of the punishment goals for both the direct and indirect measure. For each measure of punishment goals, we conducted a separate mixed model ANOVA predicting the support of the three punishment goals (retribution, special prevention, and general prevention) as within-subjects factor and the sample (pre-service and in-service teachers) as between-subjects factor. It turned out that there were only negligible differences between the groups (the detailed results of these tests, including test statistics, are available online at the OSF). Thus, we considered it reasonable to combine the two groups to one group of teachers.

Prior to testing our main hypothesis, we considered the control variables to determine whether teachers and students perceived the misbehavior itself in a similar manner. We therefore calculated a Spearman correlation between the mean ratings of all control variables, leading to a high correlation of r_s = .83. Correspondingly, differences in punishment goal preferences between teachers and students cannot be ascribed to a different perception of the student misbehavior.

Direct measure: Endorsement of punishment goals

The teachers' endorsement ratings

For the direct measure of punishment goals, as reported in Table 2, teachers indicated a generally higher endorsement of utilitarian punishment goals as compared to retribution. Specifically, special prevention was the most endorsed goal, closely followed by general prevention. Notably, retribution received substantially lower endorsement ratings from teachers.

Table 2

Means and standard deviations of the endorsement ratings of the three punishment goals

	Retribution	Special prevention	General prevention
Sample	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Teachers	3.29 (1.30)	4.42 (0.79)	4.26 (0.98)
Students	3.24 (1.62)	3.70 (1.44)	3.23 (1.74)

Note. Higher values indicate a higher endorsement of the punishment goal from the respective sample.

To statistically test this pattern, we used a repeated measures ANOVA predicting the endorsement of the three punishment goals (retribution, special prevention, and general prevention) as within-subjects factor, followed by pairwise post-hoc *t*-tests. The analysis of variance confirmed significant differences between endorsement ratings of punishment goals, $F(2, 518) = 109.20$, $p < .001$, $\hat{\eta}^2_G = .19$. As hypothesized, follow-up *t*-tests directly examining our hypothesis revealed significantly greater endorsement of special prevention compared to retribution, $t(259) = 12.61$, $p < .001$, $d = 0.73$, and general prevention, $t(259) = 2.82$, $p = .005$, $d = 0.12$, although differences were rather small in the latter comparison. Likewise, general prevention received significantly greater endorsement than retribution, $t(259) = 9.96$, $p < .001$, $d = 0.59$.

The students' endorsement ratings

As is also shown in Table 2, special prevention received the highest endorsement ratings in students. General prevention and retribution were equally endorsed punishment goals.

A repeated measures analysis of variance confirmed significant differences between endorsement ratings of punishment goals, $F(2, 474) = 12.73$, $p < .001$, $\hat{\eta}^2_G = .02$. Follow-up *t*-tests directly testing our hypothesis revealed rather unexpected results. In contrast to our hypothesis, special prevention was most endorsed and received higher endorsement than retribution, $t(237) = 5.23$, $p < .001$, $d = 0.21$, and general prevention, $t(237) = 4.17$, $p < .001$, $d = 0.20$. Differences between retribution and general prevention were negligible, $t(237) = 0.07$, $p = .943$, $d = 0.003$.

Comparing teachers' and students' endorsement ratings

Figure 6 displays the teachers' and students' endorsement of the three punishment goals showing notable differences, especially regarding the relative degree of endorsement of retribution. Specifically, whereas retribution received endorsement ratings comparable to special prevention and general prevention in students, it received substantially lower endorsement ratings compared to the other goals from teachers. To statistically test whether teachers and students actually differed in their preferences of punishment goals, we conducted a mixed model ANOVA predicting the endorsement of the three punishment goals (retribution, special prevention, and general prevention) as within-subjects factor and the sample (teachers and students) as between-subjects factor. Most importantly for our hypothesis, this analysis revealed a significant interaction of the goal to be rated and the sample, $F(2, 992) = 27.88$, $p < .001$, $\hat{\eta}^2_G = .02$.

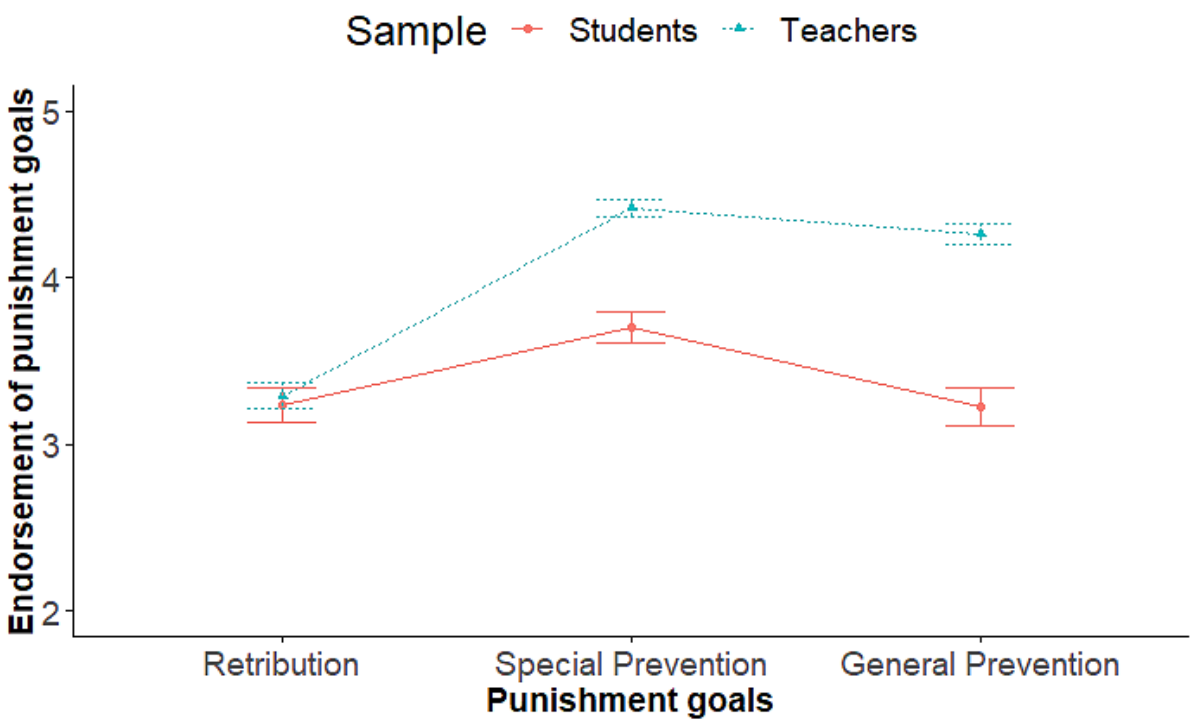

Figure 6. Results of the direct rating of punishment goals from the two samples: mean endorsement of the three punishment goals. Error bars represent one standard error of the mean.

Indirect measure: Rating of teacher punishment reactions achieving different goals

The teachers' appropriateness ratings

Descriptive statistics of the indirect measures of punishment goals are provided in Table 3. For the indirect reaction rating, contrary to the direct endorsement measure of punishment goals, teachers rated the general preventive reaction as most appropriate, closely followed by the retributive reaction. Importantly, the special preventive reaction was rated as least appropriate.

Again, we conducted a repeated measures ANOVA to compare teachers' appropriateness ratings of the three punishment reactions, followed by pairwise post-hoc *t*-tests. The analysis of variance confirmed significant differences between appropriateness ratings of the three punishment reactions, $F(2, 518) = 29.12$, $p < .001$, $\hat{\eta}^2_G = .06$. In contrast to the direct endorsement measure and contrary to our hypothesis, follow-up *t*-tests revealed that the special preventive reaction was rated significantly less appropriate than both the retributive reaction, $t(259) =$

-5.64, $p < .001$, $d = -0.34$, and the general preventive reaction, $t(259) = -8.00$, $p < .001$, $d = -0.43$. Differences between the retributive and the general preventive reaction were miniscule and not significant, $t(259) = -1.28$, $p = .202$, $d = -0.08$.

Table 3

Means and standard deviations of the ratings of the three punishment reactions

	Retributive reaction	Special preventive reaction	General preventive reaction
Sample	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Teachers	2.48 (1.57)	1.73 (1.46)	2.65 (1.50)
Students	3.84 (1.30)	3.59 (1.22)	2.85 (1.55)

Note. Higher values indicate a higher rating of the punishment reaction from the respective sample.

The students' ratings of the teachers' reactions

Prior to data analysis of the students' indirect reaction measure of punishment goals, we aggregated (that is, averaged) across the three items measuring the students' perception of each of the punishment reactions: how fair, appropriate, and just they considered each reaction to be. Cronbach's α of .90 and above (retributive reaction: $\alpha = .92$; special preventive reaction: $\alpha = .90$; general preventive reaction: $\alpha = .93$) indicated high internal consistencies of the ratings.

The students' ratings of teachers' reactions as an indirect measure of punishment goals again differed from the direct punishment goal endorsement measure. As reported in Table 3, students indicated highest ratings for the retributive reaction, closely followed by the special preventive reaction. The general preventive reaction received the lowest ratings.

Again, we conducted a repeated measures ANOVA to compare the students' perceptions of the three punishment reactions, followed by pairwise post-hoc *t*-tests. The analysis of

variance confirmed significant differences between the ratings of the three punishment reactions, $F(2, 474) = 49.42$, $p < .001$, $\hat{\eta}^2_G = .09$. Follow-up *t*-tests revealed that, as hypothesized, the retributive reaction received higher ratings than the general preventive reaction, $t(237) = 9.31$, $p < .001$, $d = 0.46$, and the special preventive reaction, $t(237) = 2.58$, $p = .010$, $d = 0.14$. However, the effect size of the latter was rather small. Furthermore, the special preventive reaction received higher ratings than the general preventive reaction, $t(237) = 7.01$, $p < .001$, $d = 0.36$.

Comparing teachers' and students' ratings of the punishment reactions

As displayed in Figure 7, teachers and students differed substantially regarding the two preventive punishment reactions. That is, both teachers and students rated the retributive reaction in a comparable manner (i.e., both perceived it as relatively appropriate). At the same time, the special preventive reaction received the lowest appropriateness ratings from teachers, whereas this reaction almost fared best in students. Likewise, the general preventive reaction yielded differences in that it received highest appropriateness ratings from teachers, but lowest ratings from students. Again, we tested whether teachers and students differed in their preferences for the three punishment reactions using a mixed model ANOVA with the three punishment reactions (retributive reaction, special preventive reaction, and general preventive reaction) as within-subjects factor and the sample (teachers and students) as between-subjects factor. Again, this analysis revealed a significant interaction of the punishment reaction and the sample, $F(2, 992) = 52.26$, $p < .001$, $\hat{\eta}^2_G = .05$.

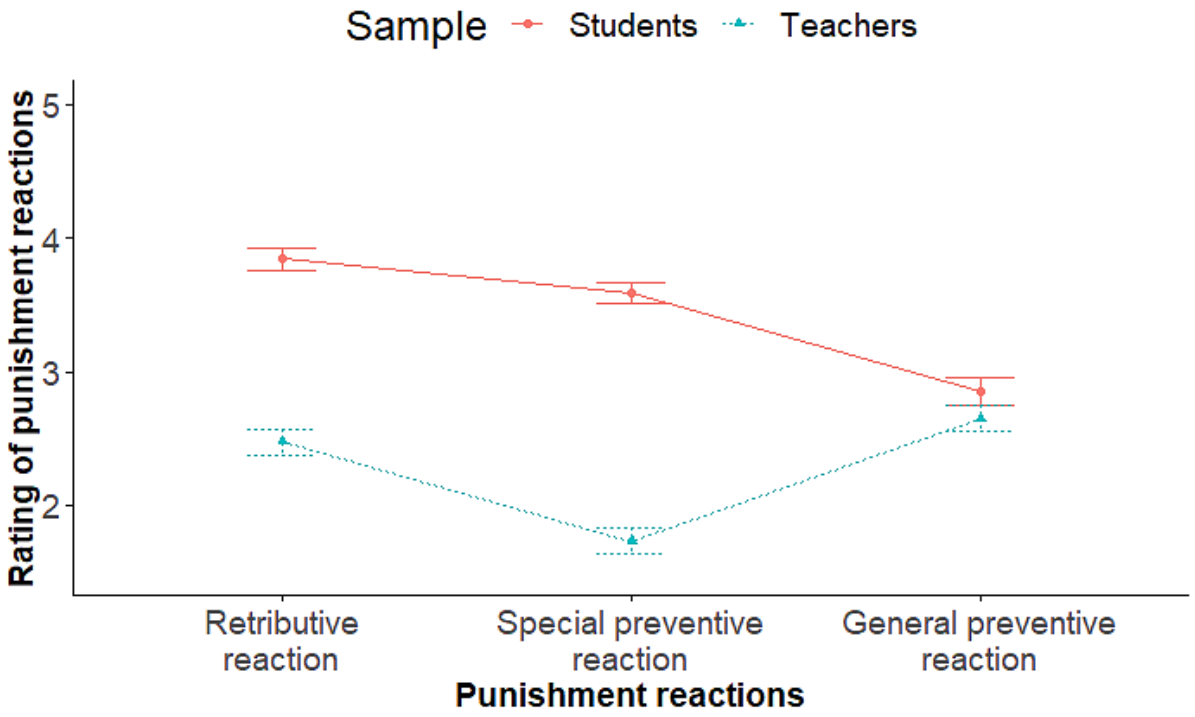

Figure 7. Results of the indirect rating measure of punishment goals from the two samples: teachers' appropriateness ratings of the three punishment reactions, and the mean of students' ratings on how fair, appropriate, and just the teachers' punishment reactions are. Error bars represent one standard error of the mean.

5.3 DISCUSSION

In everyday school life, teachers regularly have to deal with student misbehavior (Kulinna et al., 2006; Wheldall & Merrett, 1988). This is not only particularly problematic for teachers (Aloe et al., 2014; Brouwers & Tomic, 2000), but also threatens to lead to experiences of injustice in students (Fan & Chan, 1999; Israelashvili, 1997). Correspondingly, it is vital to investigate and analyze teachers' classroom intervention strategies to understand the factors enhancing students' experiences of injustice in a situation in which punishment of the misbehavior is generally justified. One key dimension of punishment behavior reflects the underlying motivation and goals of the punishment (e.g., Carlsmith et al., 2002; Gromet & Darley, 2009a). This aspect of the teachers' punishment is of particular interest, given that teachers and students

are likely to differ in their punishment goal preferences due to their unequal power positions (Mooijman et al., 2015).

In the present research, we investigated the perspectives of both teachers and school students on the purposes of punishment in a specific situation of student misbehavior. More precisely, we provided (pre-service and in-service) teachers and students with a scenario describing a student destroying the belongings of another student. We tested whether teachers and students show similar preferences in punishment goals when directly asked to indicate their endorsement of these goals. Furthermore, given that past research has revealed great differences between the explicit endorsement of punishment goals and actual punishment behavior (e.g., Carlsmith, 2008; Crockett et al., 2014), we additionally measured teachers' punishment goal preferences in a more indirect way, asking them to rate the appropriateness of different punishment reactions that are known to primarily achieve one specific goal. Importantly, we provided students with the very same punishment reactions and asked them to indicate the extent to which they perceived these reactions as fair, appropriate, and just, if shown by a teacher. In total, we thus investigated teachers' and students' preferences for retribution, special prevention, and general prevention as punishment goals and whether these preferences are comparable in teachers and students.

The main results of the present research are threefold. First, we found notable inconsistencies between the direct endorsement of punishment goals and the ratings of the reactions achieving these goals (i.e., the indirect measure of punishment goals). For example, whereas special prevention was consistently of highest preference for both teachers and students when asking for their endorsement of the goal directly, the corresponding reaction achieving special prevention received relatively low ratings. Furthermore, retribution received rather low endorsement scores (i.e., in the direct measure of punishment goals), whereas the retributive reaction received higher ratings (i.e., in the indirect measure of punishment goals). This is in line with a growing body of research showing that individuals generally endorse other goals than

they actually pursue in their punishment behavior (Applegate et al., 1996) and particularly overestimate the influence of future outcomes (i.e., utility) on their own punishment decisions (Carlsmith, 2008; Crockett et al., 2014).

Second, as hypothesized, teachers indicated a general preference for utilitarian punishment goals over retribution. This was particularly true for general prevention, whereas special prevention was only preferred over retribution in the direct measure of punishment goals. These punishment goal preferences of teachers—as individuals in a position of leadership in the school context—are consistent with research showing that individuals holding a position of power show a preference for general preventive punishment (Mooijman et al., 2015). However, the expected preference of students for retribution as the most important motivation underlying punishment was only true for their rating of the three teachers' punishment reactions (i.e., the indirect measure of punishment goals). By contrast, directly asking students for their endorsement of punishment goals, they indicated highest endorsement of special prevention, with retribution and general prevention achieving lower, but almost equivalent ratings.

Third, and most important, we found substantial differences between students' and teachers' punishment goal preferences, in particular for the support of general prevention. Whereas general prevention was consistently least supported by students both when asked to indicate their general endorsement and the rating of the teachers' punishment reactions, it received high support by teachers across measures of punishment goals. This may be particularly problematic, given that the pursuit of general prevention as the underlying motivation of punishment may have undesirable consequences. In fact, it has been shown that an authority's punishment for general preventive purposes actually leads to a decrease in rule compliance by subordinates (Mooijman et al., 2017). Correspondingly, it could be suggested that this decrease in rule compliance by subordinates is due to differing goals of people with different power positions. However, this is rather speculative and future research is needed to illuminate this further.

The present results have several theoretical and practical implications that warrant discussion. In particular, given the notable discrepancies between teachers' goals and their support of actual punishment behavior pursuing such goals, one might encourage teachers and individuals involved in teacher education to reflect on the topic of punishment in the educational setting and the motivations underlying their interactions with misbehaving students. This is especially important, given that teachers expressed a consistent preference for general prevention as the motivation underlying their punishment, whereas students' support of this goal was rather low—specifically for the general preventive punishment reaction. However, students' perception of the general preventive reaction was arguably to be expected, as one key of actual punishment behavior that is meant to prevent future misbehavior is the public display of the offender, the misbehavior, and the punishment (Carlsmith, 2006; Keller et al., 2010) and such a public reprimand has already shown to be perceived as unacceptable by students (Elliott et al., 1986).

Then again, the consistencies between teachers' and students' punishment goal preferences when directly asked for their endorsement of the goals may provide some hope that there are in fact opportunities to react to student misbehavior without triggering students' perception of injustice—and even without giving up on the goal of general prevention in punishment. Nonetheless, researchers and practitioners would need to work on generally acceptable classroom intervention strategies that achieve the preferred general preventive goal for teachers, without students inferring that punishment is carried out only to make an example of the offender and to prevent others from engaging in such misbehavior.

Furthermore, it should be mentioned that, regarding the students' view on the misbehavior and punishment, the results of the present research are limited to the role of an uninvolved observer. By contrast, in many situations of misbehavior, there are several other perspectives involved, such as a perpetrator or a victim (Schmitt et al., 2005). Therefore, future work will need to examine the students' perspective on misbehavior in the classroom and a

teacher's reaction to it from different perspectives. For example, one could suggest that a perpetrator may prefer a special preventive punishment, whereas the victim of the misbehavior may only be satisfied if the perpetrator receives a retributive punishment (i.e., the offender "gets what she deserves"). In any case, the impact of different perspectives on a teachers'—or, more generally, an authority's—punishment decision calls for further investigation in future research.

Before concluding, potential limitations of the present research should be acknowledged. In particular, we used scenarios and comic strips to investigate teachers' and students' perspectives on punishment rather than observing actual behavior in schools. Indeed, there is arguably an inherent difference between situations actually occurring in class and such prototypical scenarios (Hughes & Huby, 2004; Schoenberg & Ravdal, 2000). As a consequence, extensive care was taken to ensure that the material was suitable (e.g., by consulting teachers to evaluate and improve the material) and to increase the relevance and authenticity of the student misbehavior and the reactions of the teachers used. Additionally, similar experimental approaches have been successfully used to investigate teachers' evaluation and behavior in other domains (e.g., Baudson & Preckel, 2013). Nonetheless, future research may consider studying actual student misbehavior and punishment behavior by teachers in school settings using field observations (Klein, 2008; Klein et al., 1993; Lipshitz et al., 2001).

Further associated with the scenario and comic strip used, teachers and students were confronted with only one specific instance of student misbehavior (i.e., a student destroying the belongings of another student). Therefore, the results obtained here may be subject to unknown specifics of this scenario and the reactions offered. Indeed, it could be argued that teachers' and students' punishment goals may be influenced by other aspects of the misbehavior not addressed herein, such as the magnitude of harm caused (Carlsmith, 2006). It is up to future research to replicate and extend the present findings to more diverse forms of student misbehavior.

In conclusion, the present research is the first to compare teachers' and students' views on the purposes of punishment in the school context. In light of the findings and the observation that the approach, as a whole, is fruitful, other researchers are strongly encouraged to integrate all perspectives (i.e., the teachers' and students' views) on the psychological analysis of teaching and instruction. Finally, we hope the present findings contribute to the development of classroom intervention strategies that may reduce rather than enhance conflicts in student-teacher-interactions.

6. GENERAL DISCUSSION

The high frequency of student misbehavior and its treatment is not only a major challenge for teachers (Melnick & Meister, 2008) and a threat to their well-being (Aloe et al., 2014), but also a primary cause of experiences of injustice in students (Israelashvili, 1997). More specifically, students predominantly list incidents of the punishment of misbehavior when asked to describe a situation in school that they perceived as unjust (Fan & Chan, 1999). Therefore, it is important to examine and understand teachers' classroom intervention strategies to identify factors that may enhance students' experiences of injustice in situations in which the punishment of misbehavior is generally justified. One such factor may be the goals teachers intend to achieve when punishing student misbehavior (Carlsmith et al., 2002; Gromet & Darley, 2009a).

People generally pursue three goals when reacting to misbehavior: retribution (i.e., evening out the harm caused), special prevention (i.e., preventing recidivism of the offender), and general prevention (i.e., preventing imitation of others). Importantly, recent research suggests that people's support of these punishment goals is subject to hierarchy and power, with people in a more powerful position preferring utilitarian punishment, whereas people in a less powerful position prefer retribution (Mooijman et al., 2015).

In the school context, teachers are typically in a more powerful position than students (Goddard, 2000). Consequently, this difference in power may lead to deviations in the punishment goals teachers and students support in the treatment of student misbehavior. More precisely, students were suggested to show a preference for retributive over utilitarian punishment. Whereas, teachers were suggested to show a preference for utilitarian over retributive punishment. Eventually, such a "mismatch" in students' and teachers' punishment goals preferences may increase the extent of norm violations and, importantly, may explain students' experiences of injustice in the treatment of classroom misbehavior (Mooijman et al., 2017).

In this dissertation, I presented three projects that examined teachers' punishment and its goals along with the students' perception of classroom intervention strategies pursuing these

goals. More detailed, I first examined students' (i.e., children's) general support of different punishment goals. Furthermore, I applied an attributional approach to understand and study the goals teachers intend to achieve when reacting to student misbehavior. Finally, I investigated teachers' and students' punishment goal preferences regarding the same student misbehavior to identify potential deviations in their support of different punishment goals and reactions pursuing these goals.

On the broadest level, the results of my dissertation are fourfold. First, school students (i.e., children) revealed a preference for retribution and special prevention over general prevention as punishment goals. That is, students supported retribution and special prevention across several studies and types of measurement (projects 1 and 3). By contrast, general prevention was only supported (i.e., relative to the other goals) when children participated in an economic game (project 1) or when they were asked to explicitly indicate their endorsement of this punishment goal (project 3). For the latter, this may be due to a generally high support of utilitarian punishment for such direct punishment goal measures, as observed across all studies (see below for a more detailed discussion of this aspect). The results obtained from the economic game are more puzzling though, given that they may be attributable to several aspects. For example, one may argue that the paradigm—as a rather abstract situation—may be too artificial and far off of the natural environment of young children. Consequently, it may be too difficult for children to imagine themselves in such a situation, although there is considerable research successfully applying such approaches to developmental research questions (Gummerum et al., 2008). Furthermore, the findings may also be caused by the specific nature of the misbehavior and the punishment, as the "target" of the general preventive punishment was particularly clear in the economic game approach (i.e., communicating the punishment to a specific bystander) compared to the general preventive punishment options provided in the other situations (i.e., the scenarios). That is, in the scenarios, it was not as clear whom the "message" of the general preventive punishment would reach compared to the economic game. In total, there are several

potential explanations for the inconsistency in children's support of general preventive punishment and, with the present data, none of these explanations can be ruled out. Thus, further research is needed to investigate the conditions under which children support general prevention as a primary goal of punishment.

Second, teachers indicated a general preference for both special prevention and general prevention over retribution. This was particularly true for more direct measures of punishment goals (projects 2 and 3) and in situations in which the misbehavior was attributed to a controllable cause (project 2). Under very specific circumstances, though, teachers also demonstrated a preference for retributive over utilitarian punishment (i.e., when measured indirectly and, in particular, when the misbehavior was perceived as uncontrollable; project 2). This inconsistency, again, may be primarily due to methodological characteristics of the goal assessment, as retribution generally yielded higher support in more indirect punishment goal measures (again, see below for a more detailed discussion of this aspect). Notably, the two utilitarian punishment goals were both largely subject to the perceived controllability of the cause of the misbehavior, with a more controllable cause leading to a greater support of special and general preventive punishment. By contrast, retribution was only negligibly subject to teachers' attributions. More precisely, teachers perceived retribution as almost equally important, independent of the perceived stability or controllability of the cause of the misbehavior.

Third, these broad punishment goal preferences of students and teachers ultimately translate into a "mismatch" in their views on the legitimate purposes of punishment and the present findings suggest that this may indeed enhance students' perception of injustice (project 3). More precisely, teachers rated the general preventive punishment reaction as most appropriate, whereas students rated this reaction as least fair, appropriate, and just, compared to a retributive or special preventive reaction. However, as sketched above, there are in fact situations in which students also support general preventive punishment (project 1) and in which students' and teachers' punishment goal preferences are more similar, such as when they are explicitly

asked for their endorsement of (preventive) punishment goals (project 3). Furthermore, special prevention was identified as the generally least controversial punishment goal (i.e., both students and teachers showed rather comparable relative support of this punishment goal). In fact, these results provide some hope that students' and teachers' opinions on the purposes of punishment are not insurmountably different and that the punishment of classroom misbehavior may be possible without necessarily leading to experiences of injustice in students.

Fourth, across all projects, there were strong differences in the results obtained depending on the approaches applied to assess participants' support of the punishment goals. This was particularly apparent when comparing direct endorsement measures and rather indirect measures of punishment goals (projects 2 and 3). Specifically, the direct measures yielded generally high support of all punishment goals, with a large endorsement of utilitarian punishment, whereas retribution received less explicit endorsement. By contrast, the more indirect measure of punishment goals yielded a generally higher support of retribution, whereas utilitarian punishment received less support. This is in line with past research suggesting a difference between individuals' general endorsement of punishment goals and their actual punishment behavior in concrete situations (Carlsmith, 2008; Crockett et al., 2014).

6.1 THEORETICAL IMPLICATIONS

The results of the present research have several theoretical implications for educational research and the social psychological study of laypeople's punishment goals. First, as outlined above and in line with past research (Carlsmith, 2008; Crockett et al., 2014), the results largely depended on the approaches applied to measure participants' punishment goals. Apparently, there is an inherent difference between people's general endorsement of punishment goals (i.e., on an abstract level) and their support of actual punishment behavior achieving such goals in the concrete. Unfortunately, although these differences have been identified repeatedly in the past, it was rarely a focus in most scientific literature on laypeople's punishment goals. That is,

in most studies, scholars either directly (e.g., Giacomantonio et al., 2017; Mooijman et al., 2015) or indirectly (e.g., Keller et al., 2010) measured participants punishment goals and interpreted the results as a "true" assessment of the motivations underlying people's punishment, without considering potential effects of the specific approaches used.

However, it stands to question whether all approaches applied to assess people's punishment goals actually examine the very same construct. If yes, one may ask which the most valid way of measurement is. If no, one may ask what the two constructs actually represent and whether one of them is more important than the other. The results of the present projects corroborate past evidence (Carlsmith, 2008), suggesting two rather different representations of punishment goals that do not necessarily correlate perfectly (although the correlations of the direct and indirect measures of punishment goals obtained in the present research were generally high; e.g., see supplementary material of project 2 at the OSF). That is, as also shown in other domains of research in morality and justice-related decision making, it appears that people fundamentally differ in their abstract versus concrete ways of thinking (Nichols & Knobe, 2007). In the context of punishment, people are consistently more willing to support utilitarian purposes when thinking in the abstract, whereas they are more willing to support retribution when it comes to a concrete case of punishment.

Importantly, both types of individuals' support of different punishment goals (i.e., abstract and concrete) are of great value in their own regard. This may best be explained by a remarkable example: the *Three Strikes Initiative*. That is, during the 1994 election in California, a majority of the citizens voted for a policy, sentencing any person a life sentence that committed three felonies (i.e., pursuing a utilitarian punishment goal; Tyler & Boeckmann, 1997). However, after the laws were passed, individuals were largely unsatisfied with the policy in concrete cases of punishment, expressing their willingness to make large exceptions to the laws (Applegate et al., 1996). Thus, as this example shows, people's abstract punishment goals

(when voting for the policy) are rather crucial for policy makers. However, in the specific treatment of misbehavior (when judging specific felonies), individuals' concrete punishment goals are of importance. In sum, the present findings suggest that it is of theoretical importance to be very precise about the actual concept of interest (i.e., the support of punishment goals in the abstract vs. the support of punishment behavior achieving these goals in the concrete) and to consider the influences one may get through the measurement approaches adopted.

Second, the present research may serve as an example for educational research examining students' acceptability of classroom intervention strategies. That is, the projects outlined herein identified a potential cause of students' perceptions of injustice in school through differences between teachers and students in the motives underlying punishment. This adds to research on teachers' punishment behavior that primarily had a focus on the effectiveness of classroom management techniques in decreasing student misbehavior (e.g., Lewis et al., 2005; Little et al., 2002) and the few studies examining the students' perspective on different classroom intervention strategies (e.g., Elliott, 1986; Lewis et al., 2008).

Importantly, the present research program extends this research as all three projects were designed to identify psychological characteristics underlying the problems in the treatment of student misbehavior, rather than simply comparing teachers' and students' preferences regarding a set of specific classroom discipline practices (e.g., Lewis, 2001; Lovegrove et al., 1985). Therefore, the present research program provides a successful model that the investigation of the psychological mechanisms underlying students' and teachers' behavior may be a fruitful approach in educational research. This approach should be extended to other psychological aspects that may be of interest along these lines, for instance, the psychological needs of the people affected by an offense (Okimoto & Wenzel, 2008, 2009; Shnabel & Nadler, 2008; Shnabel, Nadler, Ullrich, Dovidio, & Carmi, 2009), which will eventually lead to a more comprehensive theoretical framework of the perception of injustice in school.

Similarly, the present research program is an example of how (justice-related) social psychological models can be applied to real world phenomenon. A related model that may also be useful for solving real world problems—and classroom conflicts and interventions in particular—is the *Needs-Based Model of Reconciliation*, as this explicitly examines the psychological needs of offenders and victims in transgressions (Shnabel & Nadler, 2008; Shnabel et al., 2009; SimanTov-Nachlieli & Shnabel, 2014). Thus, this model may be of particular interest when extending the present research to consider the students' perspectives in the roles of the misbehaving individual who is ought to be punished and the victim of the original misbehavior. Likewise, applying this model to develop fair classroom intervention strategies (i.e., applying it to a real world problem) may, in turn, contribute to the improvement of the model itself. In sum, as outlined above, misbehavior and its treatment are a fundamental challenge in the field of education and the application of psychological models to understand and improve the resolution of classroom conflicts may be very enriching, not only for practitioners but also for scholars.

6.2 PRACTICAL IMPLICATIONS

The results of the present research program also provide several practical implications, both for practitioners in school and beyond the school context. First, there were consistent differences between teachers' support of punishment goals in the abstract (i.e., when directly asked for their endorsement of the goals) and their rating of specific punishment reactions achieving different goals in the concrete. Thus, teachers should be encouraged to reflect on the goals they want to achieve when reacting to student misbehavior, as their actual, concrete punishment behavior might not be consistent with these aspired goals.

Second, students particularly expressed rather little support for general preventive punishment reactions, whereas retributive and special preventive punishment were preferred. Consequently, teachers may want to consider more retributive or special preventive punishment when reacting to student misbehavior to soothe potential conflicts with the students. Likewise,

teachers wishing to maintain the achievement of their actually preferred goal of general prevention may consider the application of classroom intervention strategies that are less likely to be perceived as particularly general preventive by students. The development and proper application of such strategies may be a notable topic in teacher education programs at universities.

Third, there were generally smaller differences between teachers' and students' punishment goal preferences in the abstract than between their preferences of concrete punishment reactions achieving these goals. In light of this finding, teachers may consider to flesh out classroom policies in collaboration with the class to manage potentially emerging student misbehavior. Receiving the students' commitment to such policies involving an abstract support of general preventive punishment may decrease the likelihood of perceived injustices, in case the policy has to be applied to treat student misbehavior (although, research on the Three Strikes Initiative suggests that such an optimistic hope may not be granted; Applegate et al., 1996).

Fourth, the present research program identified a potential problem in the "mismatch" between students' and teachers' perspectives concerning rather punitive reactions to misbehavior (i.e., the punishment was solely provided by an independently deciding punisher). However, a more recently emerged approach in the legal system is the *restorative justice philosophy* (Bazemore, 1998; Braithwaite, 1998), which considers the perspectives of the victim, the offender, and the community in which the offense occurred to assign a punishment. One key aspect of this approach is a face-to-face meeting involving all parties: the victim, the offender, and other community members (Wenzel, Okimoto, Feather, & Platow, 2010). In this meeting, the offender and the victim present their perspectives on the misbehavior and, using a consensus decision-making approach, work out an appropriate punishment for the offender with participation from all parties.

Importantly, such restorative justice procedures may resolve the otherwise existing differences in students' and teachers' views on the appropriate punishment. In fact, various schools

have already introduced justice approaches inspired by restorative justice—such as peer mediation in the case of student conflict or school community conferencing—and although most programs are still at the infancy stage, there is first evidence for its success with decreasing rates of bullying between students and more positive teacher-student-relationships (Gregory, Clawson, Davis, & Gerewitz, 2016).

However, such restorative justice approaches also entail large challenges (McCluskey et al., 2008). For example, the implementation of such restorative justice processes entail a deep change in school climate and, therefore, takes several years to run smoothly (Gregory et al., 2016). Furthermore, not all student misbehaviors can go through a comprehensive restorative justice process (Varnham, 2005). Therefore, it is nevertheless important to improve teachers' ability to independently deal with student misbehavior, but to find a punishment that is both appropriate for the teacher and perceived as fair by the students. The present findings may be helpful to achieve this in particular.

Lastly, the treatment of misbehavior is not only a challenge for teachers in the school environment, but also for authorities in other organizations (e.g., companies; Mazraeh;Arab-Khazaeli, 2013; Spector & Fox, 2010) and the legal system (Tyler, 2006). Importantly, in these environments the perceived legitimacy and fairness of punishment is as essential as in the school system (Fox, Spector, & Miles, 2001; Nadler, 2005). Indeed, there are arguably also fundamental differences in status and power between "punishers" and "observers" (or the "punished") in these contexts. Correspondingly, this may lead to a comparable "mismatch" in punishment goal preferences—and the related consequences—as observed in the present research program (Mooijman et al., 2017). Therefore, the present findings have fundamental implications for the treatment of misbehavior above and beyond the school context (Bushway & Owens, 2013).

6.3 LIMITATIONS AND FUTURE RESEARCH

There are several limitations of the present research program that have to be acknowledged. First, as outlined above, the present research is limited to the extent that it exclusively focuses on teachers and students in the roles of observers of a misbehavior. Future research may therefore extend the present findings to students in the roles of the actual individual who is ought to be punished for misbehavior or the victim of the original misbehavior. Importantly, this research should consider the specificities of such roles (e.g., their particular needs; Shnabel & Nadler, 2008).

Second, all projects focused on three specific punishment goals, namely, retribution, special prevention, and general prevention. However, several other potential goals are being discussed in the psychological and criminological literature (Keller et al., 2010; Oswald, Hupfeld, Klug, & Gabriel, 2002), such as *rehabilitation* (i.e., enabling the offender to be reintegrated into the community) or *incapacitation* (i.e., prohibiting future crimes by the offender by locking her up). Importantly, these goals are arguably of more interest for actual criminal offenses and, thus, may be less suitable for typically mild forms of student misbehaviors (Infantino & Little, 2005; Reynolds et al., 2011). Indeed, the present research focused on the punishment goals that may be most appropriate for classroom misbehavior. Future research may therefore consider other potential goals of punishment (e.g., rehabilitation and incapacitation) to receive a more comprehensive account of teachers' punishment goals. Nevertheless, given the natural overlap of rehabilitation (as a positive approach to prevent recidivism) and incapacitation (as a negative approach to prevent recidivism) with special prevention (i.e., preventing recidivism in general), results are not expected to be fundamentally different from the findings yielded in the present research.

Third, examining the students' perspective on punishment goals, the present research exclusively focused on a particular age of the children (i.e., children around the ages of 10 and 11). However, past research suggests that students from different grades differ in their fairness

evaluations of various kinds of classroom practices (Thorkildsen, 1989). Therefore, future research is needed to investigate whether the findings reported herein are applicable to students of lower or higher grades.

Fourth, most studies reported in the present work applied a scenario technique providing participants with a text (or comic) describing a specific misbehavior. Subsequently, participants were asked to indicate their feelings and thoughts regarding this scenario and potential (pretested) reactions to it. Great effort has been put into the development of this material, consulting teachers as experts to evaluate and improve the scenarios and the reactions to it. Importantly, focusing on only one scenario (i.e., one misbehavior) could be problematic as the results may be subject to the specificity of the misbehavior applied. This also relates to the punishment reactions used. These were pretested with regard to their goal achievement (i.e., achieving primarily one goal but not the other two) and their general harshness (i.e., being perceived as equally harsh). However, the reactions were arguably different in several other characteristics (e.g., entering a student's misbehavior to the class register vs. telling a misbehaving student to write an essay). Although it is unclear whether such differences had an impact on the present findings, future research may conceptually replicate and extend the present results with a more diverse set of material.

Lastly, although the present dissertation was primarily interested on an applied topic, the data in all projects were collected through questionnaires and in online environment rather than in real world settings. In particular, participants were asked to imagine a specific situation (i.e., a misbehavior) and to predict the feelings and thoughts they might have, as well as the behavior they would show when being in such a situation. However, one may expect large differences in these feelings, thoughts, and behaviors between an imagined and an actually experienced situation (Hughes & Huby, 2004; Schoenberg & Ravdal, 2000). Although the material applied in all projects was created with care and scrutiny, the inherent differences between experiences and behavior in a laboratory-like, well-structured setting and the field are likely to

persist; and this is particularly painful in light of the practical orientation of the present work. Therefore, again, the present research can only be interpreted as a first step to understand the issues associated with the treatment of student misbehavior and future research is strongly encouraged to test the patterns derived in this dissertation in actual classroom settings using observations in the field (Klein, 2008; Klein et al., 1993; Lipshitz et al., 2001).

6.4 CONCLUSIONS

The present research program analyzed and compared students' (i.e., children's) and teachers' support of different punishment goals. It has been shown that students have a preference for retribution and special prevention over general prevention as punishment goals, although general prevention was also supported under specific conditions. By contrast, teachers indicated a general preference for both special prevention and general prevention over retribution, although retribution was also supported under specific conditions. These patterns ultimately translate into a "mismatch" of teachers' and students' views on the legitimate purposes of punishment, which may, indeed, partially cause students' experiences of injustice in the treatment of classroom misbehavior. Importantly, the results obtained largely depended on the approaches applied to measure participants' punishment goals.

The present findings have several important implications, both from a theoretical and a practical point of view. Most crucial for social psychological research on punishment goals, the present results corroborate that researchers in this domain should be very clear-cut about the specific concept they are interested in, as people express fundamentally different opinions on retributive and utilitarian punishment goals when asked for their general endorsement (i.e., in the abstract), compared to their support of concrete punishment reactions achieving these goals. From a practical point of view, teachers are strongly encouraged to reflect on the goals they endorse in general and the behavior they show to achieve these goals, given the great discrepancies yielded between the teachers' evaluations of the punishment goals and the behaviors

actually achieving these goals. Furthermore, the present findings may be a first step to increase teachers' awareness of potential differences in their own punishment goals and the goals students may perceive as appropriate and fair.

However, the present research program is arguably not complete, given several limitations of the projects conducted. In particular, future research is needed to consider other important roles in situations of student misbehavior and punishment (e.g., the actual offender or victim of the misbehavior), to generalize the results to other misbehaviors (e.g., more severe misconducts), and to test the present findings in actual classroom settings in the field.

Nonetheless, the present dissertation is the first studying both teachers' and students' perspectives on the purposes of punishment in the school context. In light of the experience that the present approach—applying social psychological theories to examine real world problems—appears to be fruitful, scholars are greatly encouraged to advance research projects along these lines. Finally, the present findings may contribute to the development of classroom intervention strategies that do not increase but reduce conflicts in student-teacher-interactions.

7. REFERENCES

Aloe, A. M., Shisler, S. M., Norris, B. D., Nickerson, A. B., & Rinker, T. W. (2014). A multivariate meta-analysis of student misbehavior and teacher burnout. *Educational Research Review, 12*, 30–44. http://doi.org/10.1016/j.edurev.2014.05.003

Ambrose, M. L., Seabright, M. A., & Schminke, M. (2002). Sabotage in the workplace: The role of organizational injustice. *Organizational Behavior and Human Decision Processes, 89*, 947–965.

Andreou, E., & Rapti, A. (2010). Teachers' causal attributions for behaviour problems and perceived efficacy for class management in relation to selected interventions. *Behaviour Change, 27*(1), 53–67. http://doi.org/10.1375/bech.27.1.53

Ansari, M. E., Maleki, S. V., Mazraeh, S., & Arab-Khazaeli, H. (2013). Individual, job, and organizational predictors of Counterproductive Work Behavior. *Journal of Basic and Applied Scientific Research*, *3*(4), 78–86.

Applegate, B. K., Cullen, F. T., Turner, M. G., & Sundt, J. L. (1996). Assessing public support for three-strikes-and-you're-out laws: Global versus specific attitudes. *Crime & Delinquency, 42*(4), 517–534. http://doi.org/10.1177/0011128796042004002

Avtgis, T. A., & Rancer, A. S. (2008). The relationship between trait verbal aggressiveness and teacher burnout syndrome in K–12 teachers. *Communication Research Reports, 25*(1), 86–89. http://doi.org/10.1080/08824090701831875

Balafoutas, L., Nikiforakis, N., & Rockenbach, B. (2014). Direct and indirect punishment among strangers in the field. *Proceedings of the National Academy of Sciences, 111*(45), 15924–15927. http://doi.org/10.1073/pnas.1413170111

Baudson, T. G., & Preckel, F. (2013). Teachers' implicit personality theories about the gifted: An experimental approach. *School Psychology Quarterly, 28*(1), 37–46.

http://doi.org/10.1037/spq0000011

Bazemore, G. (1998). Restorative justice and earned redemption: Communities, victims, and offender reintegration. *American Behavioral Scientist*, *41*(6), 768–813. http://doi.org/10.1177/0002764298041006003

Ben-Ner, A., Amit, K., & Levy, O. (2008). Economic and hypothetical dictator game experiments: Incentive effects at the individual level. *The Journal of Socio-Economics*, *37*(5), 1775–1784. http://doi.org/10.1016/j.socec.2007.11.004

Bentham, J. (1962). Principles of penal law. In J. Bowring (Ed.), *'The Works of Jeremy Bentham'* (p. 396). Edinburgh: William Tait.

Blake, P. R., & McAuliffe, K. (2011). "I had so much it didn't seem fair": Eight-year-olds reject two forms of inequity. *Cognition*, *120*(2), 215–224. http://doi.org/10.1016/j.cognition.2011.04.006

Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. New York: Guilford.

Bowles, S., & Gintis, H. (2004). The evolution of strong reciprocity: Cooperation in heterogeneous populations. *Theoretical Population Biology*, *65*(1), 17–28. http://doi.org/10.1016/j.tpb.2003.07.001

Boyd, R., Gintis, H., & Bowles, S. (2010). Coordinated punishment of defectors sustains cooperation and can proliferate when rare. *Science*, *328*(5978), 617–620. http://doi.org/10.1126/science.1183665

Boyd, R., Gintis, H., Bowles, S., & Richerson, P. J. (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(6), 3531–3535. http://doi.org/10.1073/pnas.0630443100

Boyle, G. J., Borg, M. G., Falzon, J. M., & Baglioni, A. J. J. (1995). A structural model of the

dimensions of teacher stress. *British Journal of Educational Psychology*, *65*(1), 49–67.

Braithwaite, J. (1998). Restorative justice. In M. H. Tonry (Ed.), *The handbook of crime and punishment* (pp. 323–344). Oxford: Oxford University Press.

Brouwers, A., & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teaching and Teacher Education*, *16*, 239–253. http://doi.org/10.1016/S0742-051X(99)00057-8

Bru, E., Murgberg, T. A., & Stephens, P. (2001). Social support, negative life events and pupil misbehaviour among young Norwegian adolescents. *Journal of Adolescence*, *24*(6), 715–727. http://doi.org/10.1006/jado.2001.0434

Bru, E., Stephens, P., & Torsheim, T. (2002). Students' perceptions of class management and reports of their own misbehavior. *Journal of School Psychology*, *40*(4), 287–307. http://doi.org/10.1016/S0022-4405(02)00104-8

Bushway, S. D., & Owens, E. G. (2013). Framing punishment: Incarceration, recommended sentences, and recidivism. *Journal of Law and Economics*, *56*, 301–317. http://doi.org/10.1086/669715

Carlsmith, K. M. (2006). The roles of retribution and utility in determining punishment. *Journal of Experimental Social Psychology*, *42*(4), 437–451. http://doi.org/10.1016/j.jesp.2005.06.007

Carlsmith, K. M. (2008). On justifying punishment: The discrepancy between words and actions. *Social Justice Research*, *21*(2), 119–137. http://doi.org/10.1007/s11211-008-0068-x

Carlsmith, K. M., & Darley, J. M. (2008). Psychological aspects of retributive justice. *Advances in Experimental Social Psychology*, *40*(07), 193–236. http://doi.org/10.1016/S0065-2601(07)00004-4

Carlsmith, K. M., Darley, J. M., & Robinson, P. H. (2002). Why do we punish? Deterrence and just deserts as motives for punishment. *Journal of Personality and Social Psychology*, *83*(2), 284–299. http://doi.org/10.1037//0022-3514.83.2.284

Carpendale, J. (2000). Kohlberg and Piaget on stages and moral reasoning. *Developmental Review*, *20*(2), 181–205. http://doi.org/10.1006/drev.1999.0500

Chan, D. W. (1998). Stress, coping strategies, and psychological distress among secondary school teachers in Hong Kong. *American Educational Research Journal*, *35*(1), 145–163. http://doi.org/10.3102/00028312035001145

Chang, M. L. (2009). An appraisal perspective of teacher burnout: Examining the emotional work of teachers. *Educational Psychology Review*, *21*(3), 193–218. http://doi.org/10.1007/s10648-009-9106-y

Chory-Assad, R. M. (2002). Classroom justice: Perceptions of fairness as a predictor of student motivation, learning, and aggression. *Communication Quarterly*, *50*(1), 58–77. http://doi.org/10.1080/01463370209385646

Colquitt, J. A., Conlon, D. E., Wesson, M. J., & Porter, C. O. L. H. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology*, *86*(3), 425–445. http://doi.org/10.1037//0021-9010.86.3.425

Crockett, M. J., Özdemir, Y., & Fehr, E. (2014). The value of vengeance and the demand for deterrence. *Journal of Experimental Psychology: General*, *143*(6), 2279–2286. http://doi.org/10.1037/xge0000018

Cushman, F. (2015). Punishment in humans: From intuitions to institutions. *Philosophy Compass*, *10*(2), 117–133. http://doi.org/10.1111/phc3.12192

Dalbert, C., Schneidewind, U., & Saalbach, A. (2007). Justice judgments concerning grading

in school. *Contemporary Educational Psychology*, *32*(3), 420–433. http://doi.org/10.1016/j.cedpsych.2006.05.003

Dalbert, C., & Stoeber, J. (2005). The belief in a just world and distress at school. *Social Psychology of Education*, *8*(2), 123–135. http://doi.org/10.1007/s11218-005-1835-2

Darley, J. M., Carlsmith, K. M., & Robinson, P. H. (2000). Incapacitation and just deserts as motives for punishment. *Law and Human Behavior*, *24*(6), 659–683. http://doi.org/10.1023/A:1005552203727

Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R., & Smirnov, O. (2007). Egalitarian motives in humans. *Nature*, *446*(7137), 794–796. http://doi.org/10.1038/nature05651

Donat, M., Gallschütz, C., & Dalbert, C. (2018). The relation between students' justice experiences and their school refusal behavior. *Social Psychology of Education*, *21*(2), 447–475. http://doi.org/10.1007/s11218-017-9423-9

Donat, M., Knigge, M., & Dalbert, C. (2018). Being a good or a just teacher: Which experiences of teachers' behavior can be more predictive of school bullying? *Aggressive Behavior*, *44*(1), 29–39. http://doi.org/10.1002/ab.21721

Donat, M., Umlauft, S., Dalbert, C., & Kamble, S. V. (2012). Belief in a just world, teacher justice, and bullying behavior. *Aggressive Behavior*, *38*(3), 185–93. http://doi.org/10.1002/ab.21421

Elliott, S. N. (1986). Children's ratings of the acceptability of classroom interventions for misbehavior: Findings and methodological considerations. *Journal of School Psychology*, *24*(1), 23–35. http://doi.org/10.1016/0022-4405(86)90039-7

Elliott, S. N. (1988). Acceptability of behavioral treatments: Review of variables that influence treatment selection. *Professional Psychology: Research and Practice*, *19*(1), 68–80.

Elliott, S. N., Witt, J. C., Galvin, G. A., & Moe, G. L. (1986). Children's involvement in intervention selection: Acceptability of interventions for misbehaving peers. *Professional Psychology, Research and Practice*, *17*(3), 235–241. http://doi.org/10.1037/0735-7028.17.3.235

Engel, C. (2011). Dictator games: A meta study. *Experimental Economics*, *14*(4), 583–610. http://doi.org/10.1007/s10683-011-9283-7

Ertesvåg, S. K., & Vaaland, G. S. (2007). Prevention and reduction of behavioural problems in school: An evaluation of the Respect program. *Educational Psychology*, *27*(6), 713–736. http://doi.org/10.1080/01443410701309258

Fan, R. M., & Chan, S. C. N. (1999). Students' perceptions of just and unjust experiences in school. *Educational and Child Psychology*, *16*(2), 32–50.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. http://doi.org/10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. http://doi.org/10.3758/BF03193146

Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature*, *454*(7208), 1079–1083. http://doi.org/10.1038/nature07155

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, *25*(2), 63–87. http://doi.org/10.1016/S1090-5138(04)00005-4

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*(6868), 137–40. http://doi.org/10.1038/415137a

FeldmanHall, O., Sokol-Hessner, P., Van Bavel, J. J., & Phelps, E. A. (2014). Fairness

violations elicit greater punishment on behalf of another than for oneself. *Nature Communications*, *5*(9), 5306. http://doi.org/10.1038/ncomms6306

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. East Norwalk: Appleton-Century-Crofts.

Finn, J. D., Fish, R. M., & Scott, L. A. (2008). Educational sequelae of high school misbehavior. *Journal of Educational Research*, *101*(5), 259–274. http://doi.org/10.3200/JOER.101.5.259-274

Fox, S., Spector, P. E., & Miles, D. (2001). Counterproductive Work Behavior (CWB) in response to job stressors and organizational justice: Some mediator and moderator tests for autonomy and emotions. *Journal of Vocational Behavior*, *59*(3), 291–309. http://doi.org/10.1006/jvbe.2001.1803

Furnham, A., & Siegel, E. M. (2011). Reactions to organizational injustice: Counter work behaviors and the insider threat. In E. Kals & J. Maes (Eds.), *Justice and Conflicts* (pp. 199–217). Berlin, Heidelberg: Springer.

Gardner, A., & West, S. A. (2004). Cooperation and punishment, especially in humans. *The American Naturalist*, *164*(6), 753–764. http://doi.org/10.1086/425623

Gerber, M. M., & Jackson, J. (2013). Retribution as revenge and retribution as just deserts. *Social Justice Research*, *26*(1), 61–80. http://doi.org/10.1007/s11211-012-0174-7

Gerlach, T. M., Allemand, M., Agroskin, D., & Denissen, J. J. A. (2012). Justice sensitivity and forgiveness in close interpersonal relationships: The mediating role of mistrustful, legitimizing, and pro-relationship cognitions. *Journal of Personality*, *80*(5), 1373–1413. http://doi.org/10.1111/j.1467-6494.2012.00762.x

Giacomantonio, M., & Pierro, A. (2014). Individual differences underlying punishment motivation. The role of need for cognitive closure. *Social Psychology*, *45*(6), 449–457.

http://doi.org/10.1027/1864-9335/a000211

Giacomantonio, M., Pierro, A., Baldner, C., & Kruglanski, A. (2017). Need for closure, torture, and punishment motivations: The mediating role of moral foundations. *Social Psychology, 48*(6), 335–347. http://doi.org/10.1027/1864-9335/a000321

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior*, *24*(3), 153–172. http://doi.org/10.1016/S1090-5138(02)00157-5

Goddard, J. T. (2000). Teaching in turbulent times: Teachers' perceptions of the effects of external factors on their professional lives. *Alberta Journal of Educational Research, 46*(4), 293-310.

Goodwin, G. P., & Benforado, A. (2015). Judging the goring ox: Retribution directed toward animals. *Cognitive Science*, *39*(3), 619–646. http://doi.org/10.1111/cogs.12175

Goodwin, G. P., & Gromet, D. M. (2014). Punishment. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(5), 561–572. http://doi.org/10.1002/wcs.1301

Graham, S., Weiner, B., & Zucker, G. S. (1997). An attributional analysis of punishment goals and public reactions to O. J. Simpson. *Personality and Social Psychology Bulletin, 23*(4), 331–346. http://doi.org/0803973233

Gregory, A., Clawson, K., Davis, A., & Gerewitz, J. (2016). The promise of restorative practices to transform teacher-student relationships and achieve equity in school discipline. *Journal of Educational and Psychological Consultation*, *26*(4), 325–353. http://doi.org/10.1080/10474412.2014.929950

Gromet, D. M., & Darley, J. M. (2009a). Punishment and beyond: Achieving justice through the satisfaction of multiple goals. *Law and Society Review*, *43*(1), 1–38. http://doi.org/10.1111/j.1540-5893.2009.00365.x

Gromet, D. M., & Darley, J. M. (2009b). Retributive and restorative justice: Importance of crime severity and shared identity in people's justice responses. *Australian Journal of Psychology*, *61*(1), 50–57. http://doi.org/10.1080/00049530802607662

Gummerum, M., & Chu, M. T. (2014). Outcomes and intentions in children's, adolescents', and adults' second- and third-party punishment behavior. *Cognition*, *133*(1), 97–103. http://doi.org/10.1016/j.cognition.2014.06.001

Gummerum, M., Hanoch, Y., & Keller, M. (2008). When child development meets economic game theory: An interdisciplinary approach to investigating social development. *Human Development*, *51*, 235–261. http://doi.org/10.1159/000151494

Hainmüller, B. (2006). Exploring values in teacher education - strategies from a European partnership. In M. Persson (Ed.), *A vision of European teaching and learning - perspectives on the new role of the teacher* (pp. 115–130). Karlstad, Sweden: The Learning Teacher Network.

Hamann, K., Warneken, F., Greenberg, J. R., & Tomasello, M. (2011). Collaboration encourages equal sharing in children but not in chimpanzees. *Nature*, *476*(7360), 328–331. http://doi.org/10.1038/nature10278

Hamlin, J. K., Wynn, K., Bloom, P., & Mahajan, N. (2011). How infants and toddlers react to antisocial others. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(50), 19931–6. http://doi.org/10.1073/pnas.1110306108

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., … Ziker, J. (2006). Costly punishment across human societies. *Science*, *312*, 1767–1770. http://doi.org/10.1126/science.1127333

Hertwig, R., & Ortmann, A. (2008). Deception in social psychological experiments: Two misconceptions and a research agenda*. *Social Psychology Quarterly*, *71*(3), 222–227. http://doi.org/10.1177/019027250807100304

Holtz, B. C., & Harold, C. M. (2013). Interpersonal justice and deviance: The moderating effects of interpersonal justice values and justice orientation. *Journal of Management*, *39*(2), 339–365. http://doi.org/10.1177/0149206310390049

Hughes, R. (1998). Considering the vignette technique and its application to a study of drug injecting and HIV risk and safer behaviour. *Sociology of Health & Illness*, *20*(3), 381–400. http://doi.org/10.1111/1467-9566.00107

Hughes, R., & Huby, M. (2004). The construction and interpretation of vignettes in social research. *Social Work and Social Sciences Review*, *11*(1), 36–51. http://doi.org/10.1921/17466105.11.1.36

Infantino, J., & Little, E. (2005). Students' perceptions of classroom behaviour problems and the effectiveness of different disciplinary methods. *Educational Psychology*, *25*(5), 491–508. http://doi.org/10.1080/01443410500046549

Ingersoll, R. M. (2001). Teacher turnover, teacher shortages, and the organization of schools. *American Educational Research Journal*, *38*(3), 499–534.

Israelashvili, M. (1997). Situational determinants of school students' feelings of injustice. *Elementary School Guidance & Counseling*, *31*(4), 283–292.

Jensen, K., Call, J., & Tomasello, M. (2007). Chimpanzees are vengeful but not spiteful. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(32), 13046–13050. http://doi.org/10.1073/pnas.0705555104

Jordan, J. J., McAuliffe, K., & Warneken, F. (2014). Development of in-group favoritism in children's third-party punishment of selfishness. *Proceedings of the National Academy of Sciences*, *111*(35), 12710–12715. http://doi.org/10.1073/pnas.1402280111

Kamble, S. V, & Dalbert, C. (2012). Belief in a just world and wellbeing in Indian schools. *International Journal of Psychology*, *47*(4), 269–78.

http://doi.org/10.1080/00207594.2011.626047

Kant, I. (1952). The science of right (W. Hastie, trans.). In R. M. Hutchins & M. J. Adler (Eds.), *Great Books of the Western World* (pp. 397–446). Edinburgh: Encyclopedia Britannica.

Keller, L. B., Oswald, M. E., Stucki, I., & Gollwitzer, M. (2010). A closer look at an eye for an eye: Laypersons' punishment decisions are primarily driven by retributive motives. *Social Justice Research*, *23*(2), 99–116. http://doi.org/10.1007/s11211-010-0113-4

Kenward, B., & Östh, T. (2012). Enactment of third-party punishment by 4-year-olds. *Frontiers in Psychology*, *3*, 1–9. http://doi.org/10.3389/fpsyg.2012.00373

Killen, M., & Smetana, J. G. (2014). *Handbook of Moral Development* (2nd ed.). New York, NY: Psychology Press.

Killen, M., & Smetana, J. G. (2015). Origins and development of morality. In M. E. Lamb (Ed.), *Handbook of child psychology and developmental science* (Vol. 3, pp. 701–749). New York: Wiley-Blackwell.

Klein, G. A. (2008). Naturalistic Decision Making. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 456–460. http://doi.org/10.1518/001872008X288385

Klein, G. A., Orasanu, J., Calderwood, R., & Zsambok, C. E. (Eds.) (1993). *Decision making in action: Models and methods*. Westport, CT: Ablex Publishing.

Kohlberg, L. (1981). *Essays in moral development: The philosophy of moral development*. San Francisco: Harper & Row.

Kulinna, P. H. (2008). Teachers' attributions and strategies for student misbehavior. *Journal of Classroom Interaction*, *42*, 21–30.

Kulinna, P. H., Cothran, D. J., & Regualos, R. (2006). Teachers' reports of student

misbehavior in physical education. *Research Quarterly for Exercise and Sport*, *77*(1), 32–40. http://doi.org/10.1080/02701367.2006.10599329

Landeskriminalamt Nordrhein-Westfalen. (2017). *Polizeiliche Kriminalstatistik Nordrhein-Westfalen 2017*. Retrieved from https://polizei.nrw/artikel/polizeiliche-kriminalstatistik-2017-1

Lehmann, E. L. (2006). *Nonparametrics: Statistical methods based on ranks*. New York, NY: Springer.

Lerner, M. J. (1977). The justice motive: Some hypotheses as to its origins and forms. *Journal of Personality*, *45*(1), 1–52. http://doi.org/10.1111/j.1467-6494.1977.tb00591.x

Lerner, M. J. (1980). *The belief in a just world: A fundamental delusion*. New York, NY: Plenum Press.

Lewis, R. (1999). Teachers coping with the stress of classroom discipline. *Social Psychology of Education*, *3*, 155–171. http://doi.org/10.1023/A:1009627827937

Lewis, R. (2001). Classroom discipline and student responsibility: The students' view. *Teaching and Teacher Education*, *17*(3), 307–319. http://doi.org/10.1016/S0742-051X(00)00059-7

Lewis, R., Romi, S., Katz, Y. J., & Qui, X. (2008). Students' reaction to classroom discipline in Australia, Israel, and China. *Teaching and Teacher Education*, *24*(3), 715–724. http://doi.org/10.1016/j.tate.2007.05.003

Lewis, R., Romi, S., Qui, X., & Katz, Y. J. (2005). Teachers' classroom discipline in Australia, China and Israel. *Teaching and Teacher Education*, *21*, 729–741.

Lipshitz, R., Klein, G., Orasanu, J., & Salas, E. (2001). Taking stock of naturalistic decision making. *Journal of Behavioral Decision Making*, *14*(5), 331–352. http://doi.org/10.1002/bdm.381

Little, E. (2005). Secondary school teachers' perceptions of students' problem behaviours. *Educational Psychology*, *25*(4), 369–377. http://doi.org/10.1080/01443410500041516

Little, E., Hudson, A., & Wilks, R. (2002). The efficacy of written teacher advice (tip sheets) for managing classroom behaviour problems. *Educational Psychology*, *22*(3), 251–266. http://doi.org/10.1080/01443410220138494

Lobue, V., Nishida, T., Chiong, C., Deloache, J. S., & Haidt, J. (2011). When getting something good is bad: Even three-year-olds react to inequality. *Social Development*, *20*(1), 154–170. http://doi.org/10.1111/j.1467-9507.2009.00560.x

Lovegrove, M. N., Lewis, R., Fall, C., & Lovegrove, H. (1985). Students' preferences for discipline practices in schools. *Teaching and Teacher Education*, *1*(4), 325–333. http://doi.org/10.1016/0742-051X(85)90021-6

McAuliffe, K., Blake, P. R., & Warneken, F. (2014). Children reject inequity out of spite. *Biology Letters*, *10*, 20140743. http://doi.org/10.1098/rsbl.2014.0743

McAuliffe, K., Jordan, J. J., & Warneken, F. (2015). Costly third-party punishment in young children. *Cognition*, *134*, 1–10. http://doi.org/10.1016/j.cognition.2014.08.013

McCluskey, G., Lloyd, G., Kane, J., Riddell, S., Stead, J., & Weedon, E. (2008). Can restorative practices in schools make a difference? *Educational Review*, *60*(4), 405–417. http://doi.org/10.1080/00131910802393456

McCormick, J., & Barnett, K. (2011). Teachers' attributions for stress and their relationships with burnout. *International Journal of Educational Management*, *25*(3), 278–293. http://doi.org/10.1108/09513541111120114

McCullough, M. E., Kurzban, R., & Tabak, B. A. (2013). Cognitive systems for revenge and forgiveness. *Behavioral and Brain Sciences*, *36*(01), 1–15. http://doi.org/10.1017/S0140525X11002160

Melnick, S. A., & Meister, D. G. (2008). A comparison of beginning and experienced teachers' concerns. *Educational Research Quarterly*, *31*(3), 40–56. http://doi.org/10.1017/CBO9781107415324.004

Mooijman, M., van Dijk, W. W., Ellemers, N., & van Dijk, E. (2015). Why leaders punish: A power perspective. *Journal of Personality and Social Psychology*, *109*(1), 75–89. http://doi.org/10.1037/pspi0000021

Mooijman, M., van Dijk, W. W., van Dijk, E., & Ellemers, N. (2017). On sanction-goal justifications: How and why deterrence justifications undermine rule compliance. *Journal of Personality and Social Psychology*, *112*(4), 577–588. http://doi.org/10.1037/pspi0000084

Nadler, J. (2005). Flouting the law. *Texas Law Review*, *83*, 1399–1441.

Nichols, S., & Knobe, J. (2007). Moral responsibility and determinism: The cognitive science of folk intuitions. *Nous*, *41*(4), 663–685. http://doi.org/10.1111/j.1468-0068.2007.00666.x

Okimoto, T. G., & Wenzel, M. (2008). The symbolic meaning of transgressions: Towards a unifying framework of justice restoration. In K. A. Hegtvedt & J. Clay-Warner (Eds.), *Advances in group processes* (Vol. 25, pp. 291–326). http://doi.org/10.1016/S0882-6145(08)25004-6

Okimoto, T. G., & Wenzel, M. (2009). Punishment as restoration of group and offender values following a transgression: Value consensus through symbolic labelling and offender reform. *European Journal of Social Psychology*, *39*(3), 346–367. http://doi.org/10.1002/ejsp.537

Orth, U. (2003). Punishment goals of crime victims. *Law and Human Behavior*, *27*(2), 173–186. http://doi.org/10.1023/A:1022547213760

Ortmann, A., & Hertwig, R. (2002). The costs of deception: Evidence from psychology. *Experimental Economics*, *5*(2), 111–131. http://doi.org/10.2139/ssrn.317861

Oswald, M. E., Hupfeld, J., Klug, S. C., & Gabriel, U. (2002). Lay-perspectives on criminal deviance, goals of punishment, and punitivity. *Social Justice Research*, *15*(2), 85–98. http://doi.org/10.1023/A:1019928721720

Peter, F., & Dalbert, C. (2010). Do my teachers treat me justly? Implications of students' justice experience for class climate experience. *Contemporary Educational Psychology*, *35*(4), 297–305. http://doi.org/10.1016/j.cedpsych.2010.06.001

Peter, F., Kloeckner, N., Dalbert, C., & Radant, M. (2012). Belief in a just world, teacher justice, and student achievement: A multilevel study. *Learning and Individual Differences*, *22*(1), 55–63. http://doi.org/10.1016/j.lindif.2011.09.011

Pretsch, J., & Ehrhardt-Madapathi, N. (2018). Experiences of justice in school and attitudes towards democracy: A matter of social exchange? *Social Psychology of Education*, *21*(3), 655–675. http://doi.org/10.1007/s11218-018-9435-0

Ratcliff, N. J., Jones, C. R., Costner, R. H., Savage-Davis, E., & Hunt, G. H. (2010). The elephant in the classroom: The impact of misbehavior on classroom climate. *Education*, *131*(2), 306–314.

Resh, N., & Sabbagh, C. (2014). Sense of justice in school and civic attitudes. *Social Psychology of Education*, *17*(1), 51–72. http://doi.org/10.1007/s11218-013-9240-8

Reyna, C., & Weiner, B. (2001). Justice and utility in the classroom: An attributional analysis of the goals of teachers' punishment and intervention strategies. *Journal of Educational Psychology*, *93*(2), 309–319. http://doi.org/10.1037/0022-0663.93.2.309

Reynolds, K., Stephenson, J., & Beaman, R. (2011). Teacher perceptions of non-compliance in rural primary schools in New South Wales. *Education in Rural Australia*, *21*(2), 105–

124.

Riedl, K., Jensen, K., Call, J., & Tomasello, M. (2015). Restorative justice in children. *Current Biology*, *25*(13), 1731–1735. http://doi.org/10.1016/j.cub.2015.05.014

Ringer, M. M., Doerr, P. F., Hollenshead, J. H., & Wills, G. D. (1993). Behavior problems in the classroom: A national survey of inter ventions used by classroom teachers. *Psychology in the Schools*, *30*(2), 168–175. http://doi.org/10.1002/1520-6807(199304)30:2<168::AID-PITS2310300209>3.0.CO;2-2

Robbins, E., & Rochat, P. (2011). Emerging signs of strong reciprocity in human ontogeny. *Frontiers in Psychology*, *2*, 353. http://doi.org/10.3389/fpsyg.2011.00353

Rucker, D. D., Polifroni, M., Tetlock, P. E., & Scott, A. L. (2004). On the assignment of punishment: The impact of general-societal threat and the moderating role of severity. *Personality and Social Psychology Bulletin*, *30*(6), 673–684. http://doi.org/10.1177/0146167203262849

Sabbagh, C., & Resh, N. (2016). Unfolding justice research in the realm of education. *Social Justice Research*, *29*(1), 1–13. http://doi.org/10.1007/s11211-016-0262-1

Schmidt, M. F. H., Rakoczy, H., & Tomasello, M. (2012). Young children enforce social norms selectively depending on the violator's group affiliation. *Cognition*, *124*(3), 325–333. http://doi.org/10.1016/j.cognition.2012.06.004

Schmidt, M. F. H., & Sommerville, J. A. (2011). Fairness expectations and altruistic sharing in 15-month-old human infants. *PLoS ONE*, *6*(10), e23223. http://doi.org/10.1371/journal.pone.0023223

Schmitt, M., & Dörfel, M. (1999). Procedural injustice at work, justice sensitivity, job satisfaction and psychosomatic well-being. *European Journal of Social Psychology*, *29*(4), 443–453. http://doi.org/10.1002/(SICI)1099-0992(199906)29:4<443::AID-

EJSP935>3.0.CO;2-C

Schmitt, M., Gollwitzer, M., Maes, J., & Arbach, D. (2005). Justice sensitivity. *European Journal of Psychological Assessment*, *21*(3), 202–211. http://doi.org/10.1027/1015-5759.21.3.202

Schoenberg, N. E., & Ravdal, H. (2000). Using vignettes in awareness and attitudinal research. *International Journal of Social Research Methodology*, *3*(1), 63–74. http://doi.org/10.1080/136455700294932

Shaw, A., & Olson, K. R. (2012). Children discard a resource to avoid inequity. *Journal of Experimental Psychology: General*, *141*(2), 382–395. http://doi.org/10.1037/a0025907

Shnabel, N., & Nadler, A. (2008). A needs-based model of reconciliation: Satisfying the differential emotional needs of victim and perpetrator as a key to promoting reconciliation. *Journal of Personality and Social Psychology*, *94*(1), 116–132. http://doi.org/10.1037/0022-3514.94.1.116

Shnabel, N., Nadler, A., Ullrich, J., Dovidio, J. F., & Carmi, D. (2009). Promoting reconciliation through the satisfaction of the emotional needs of cictimized and perpetrating group members: The needs-based model of reconciliation. *Personality and Social Psychology Bulletin*, *35*(8), 1021–1030. http://doi.org/10.1177/0146167209336610

SimanTov-Nachlieli, I., & Shnabel, N. (2014). Feeling both victim and perpetrator investigating duality within the needs-based model. *Personality and Social Psychology Bulletin*, *40*(3), 301–314. http://doi.org/10.1177/0146167213510746

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 word solution. *Dialogue: The Official Newsletter of the Society for Personality and Social Psychology*, *26*(2), 4–7. Retrieved from http://www.ssrn.com/abstract=2160588

Skinner, B. F. (1938). *The behavior of organisms*. Cambridge, Massachusetts: B.F. Skinner Foundation.

Smith, C. E., Blake, P. R., & Harris, P. L. (2013). I should but I won't: Why young children endorse norms of fair sharing but do not follow them. *PLoS ONE*, *8*(3). http://doi.org/10.1371/journal.pone.0059510

So geht es an einer Grundschule in Sachsen-Anhalt zu. (2018). Retrieved from https://www.welt.de/politik/deutschland/article173881654/In-Sachsen-Anhalt-verzweifelt-eine-Grundschule-an-ihren-Schuelern.html

Soodak, L. C., & Podell, D. M. (1994). Teachers' thinking about difficult-to-teach students. *The Journal of Educational Research*, *88*(1), 44–51. http://doi.org/10.1080/00220671.1994.9944833

Spector, P. E., & Fox, S. (2010). Counterproductive Work Behavior and Organisational Citizenship Behavior: Are they opposite forms of active behavior? *Applied Psychology*, *59*(1), 21–39. http://doi.org/10.1111/j.1464-0597.2009.00414.x

Sugai, G., & Horner, R. (2006). A promising approach for expanding and sustaining school-wide positive behavior support. *School Psychology Review*, *35*(2), 245–259.

Tetlock, P. E. (2002). Social functionalist frameworks for judgment and choice: Intuitive politicians, theologians, and prosecutors. *Psychological Review*, *109*(3), 451–471. http://doi.org/10.1037//0033-295X.109.3.451

Thielmann, I., Heck, D. W., & Hilbig, B. E. (2016). Anonymity and incentives: An investigation of techniques to reduce socially desirable responding in the Trust Game. *Judgment and Decision Making*, *11*(5), 527–536.

Thorkildsen, T. A. (1989). Justice in the classroom: The student's view. *Child Development*, *60*(2), 323–334.

Tisak, M. S. (1986). Children's conceptions of parental authority. *Child Development*, *57*(August 1984), 166–176. http://doi.org/10.2307/1130648

Tollefson, N. (2000). Classroom applications of cognitive theories of motivation. *Educational Psychology Review*, *12*(1), 63–83. http://doi.org/10.1023/A:1009085017100

Turco, T. L., & Elliott, S. N. (1986). Assessment of students' acceptability ratings of teacher-initiated interventions for classroom misbehavior. *Journal of School Psychology*, *24*(3), 277–283. http://doi.org/10.1016/0022-4405(86)90060-9

Tyler, T. R. (2006). Psychological perspectives on legitimacy and legitimation. *Annual Review of Psychology*, *57*(1), 375–400. http://doi.org/10.1146/annurev.psych.57.102904.190038

Tyler, T. R., & Boeckmann, R. J. (1997). Three strikes and you are out, but why? The psychology of public support for punishing rule breakers. *Law and Society Review*, *31*(2), 237–266.

Uhlmann, E. L., Pizarro, D. A., & Diermeier, D. (2015). A person-centered approach to moral judgment. *Perspectives on Psychological Science*, *10*(1), 72–81. http://doi.org/10.1177/1745691614556679

Varnham, S. (2005). Seeing things differently: Restorative justice and school discipline. *Education and the Law*, *17*(3), 87–104. http://doi.org/10.1080/09539960500334061

Von Brock, M. B., & Elliott, S. N. (1987). Influence of treatment effectiveness information on the acceptability of classroom interventions. *Journal of School Psychology*, *25*(2), 131–144. http://doi.org/10.1016/0022-4405(87)90022-7

Wald, J., & Losen, D. J. (2003). Defining and redirecting a school-to-prison pipeline. *New Directions for Youth Development*, (99), 9–15. http://doi.org/10.1002/yd.51

Weiner, B. (2003). The classroom as a courtroom. *Social Psychology of Education*, *6*(1), 3–

15. http://doi.org/10.1023/A:1021736217048

Weiner, B., Graham, S., & Reyna, C. (1997). An attributional examination of retributive versus utilitarian philosophies of punishment. *Social Justice Research*, *10*(4), 431–452. http://doi.org/10.1007/BF02683293

Wenzel, M., Lawrence-Wood, E., Okimoto, T. G., & Hornsey, M. J. (2017). A long time coming: Delays in collective apologies and their effects on sincerity and forgiveness. *Political Psychology*, *39*(3). http://doi.org/10.1111/pops.12421

Wenzel, M., Okimoto, T. G., Feather, N. T., & Platow, M. J. (2010). Justice through consensus: Shared identity and the preference for a restorative notion of justice. *European Journal of Social Psychology*, *40*(6), 909–930. http://doi.org/10.1002/ejsp.657

Wenzel, M., Okimoto, T. G., Hornsey, M. J., Lawrence-Wood, E., & Coughlin, A. M. (2017). The mandate of the collective: Apology representativeness determines perceived sincerity and forgiveness in intergroup contexts. *Personality and Social Psychology Bulletin*, *43*(6), 758–771. http://doi.org/10.1177/0146167217697093

Wheldall, K., & Merrett, F. (1988). Which classroom behaviours do primary school teachers say they find most troublesome? *Educational Review*, *40*(1), 13–27. http://doi.org/10.1080/0013191880400102

ACKNOWLEDGMENTS

This dissertation would not have been possible without the support of numerous people whom I wish to thank wholeheartedly.

First, I would like to express my sincere gratitude to my advisor and mentor Benjamin Hilbig. He gave me great freedom to develop a research program I am truly interested in, but provided guidance whenever needed. I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my questions and queries so promptly.

I would also like to express special thanks to Manfred Schmitt for inspiring discussions and extremely valuable comments on my research. On various occasions, he willingly provided helpful feedback and improved my work with his immense enthusiasm and knowledge.

Further, I would like to thank the team of the graduate school, in particular, Regina Hölzmann and Heidrun Ludwig, and all of my colleagues who took this journey with me. Above all, thanks to my office mate and friend Svenja Matheis, for continuously keeping me updated, for always lending a sympathetic ear, and—most importantly—for sharing our secret cupboard. This research would also not have been possible without the impressive work of my great student research assistants: Fritzi, Manu, Anja, Matthias, Tanita, and Janna – thank you very much!

Moreover, I would like to sincerely thank my colleagues and friends from the Cognitive Psychology Lab Landau and the KMPTNZ without whom the past three years would have been only half as much fun. This team is a great model for a perfect mix of working with scientific rigor while having a good time (note: that's the politically correct description). Particularly, I would like to express my very great appreciation to Isabel Thielmann und Sina Klein for tremendously helpful comments on previous versions of this thesis.

Finally, I would like to express my heartfelt gratitude to my friends and family. My brothers, Torsten and Steffen, who both had great impact on my personal development during childhood and beyond. Steffen and his little family—as well as my grandparents Anna, Viktor, Werner, and Ruth—always remind me of what really counts in life. Torsten and his fiancée Nadja are the best role models one can have to set goals high, being an impressive example of dedication and perseverance in all aspects of personal and professional life. I am further deeply indebted to my parents, Birgit and Joachim, for their endless support, their unlimited loyalty, and for providing me with the certainty that I can always rely on them. They both are true everyday heroes I look up to day by day.

My deepest thanks lastly go to my girlfriend and soulmate Marie-Sophie for her patience, invaluable support, and unconditional love—not only during the last challenging months, but throughout the past ten years. There are no words to describe how much I owe her.

STATEMENT OF ORIGINALITY

I hereby declare that I have written the present dissertation independently, without assistance from external parties and without use of other resources than those indicated. The ideas taken directly or indirectly from external sources (including electronic sources) are duly acknowledged in the text. This thesis, either in full or in part, has not been previously submitted for grading at this or any other academic institution.

Heidelberg, November 2018

Mathias Twardawski

Address: Rohrbacher Straße 168 – 69126 Heidelberg

Phone: +49 176-78779144 – **E-Mail:** mathias.twardawski@gmail.com

CURRICULUM VITAE

ACADEMICS

Since Nov 2018	Research Associate in the Social Psychology Lab, Prof. Mario Gollwitzer, Ludwig-Maximilians-Universität München, Munich, Germany.
Nov 2015 – Nov 2018	Doctoral Student in the Graduate School "Teaching and Learning Processes", University of Koblenz-Landau, Germany. Research Associate in the Cognitive Psychology Lab, Prof. Benjamin E. Hilbig, University of Koblenz-Landau, Landau, Germany.
Aug 2015 – Oct 2015	Student Research Assistant in the Graduate School "Teaching and Learning Processes", University of Koblenz-Landau, Landau, Germany.
Mar 2013 – Jun 2013	Research Intern in the Social Perception and Evaluation Lab, Prof. Jay van Bavel, New York University, New York, USA.
Mar 2013 – Mar 2014	Student Research Assistant in the Political Psychology Lab – Institute for Communication and Media Psychology, Prof. Tobias Rothmund, University of Koblenz-Landau, Landau, Germany.
Dec 2010 – Jun 2014	Student Research Assistant in the Clinical Psychology Lab – Ambulance for Psychological Therapy (WiPP), Prof. Annette Schröder, University of Koblenz-Landau, Landau, Germany.

EDUCATION

Sep 2013 – Oct 2015	Master of Science in Psychology, University of Mannheim, Germany.
Jan 2015 – May 2015	Master of Science in Psychology, University of Jyväskylä, Finnland.
Oct 2009 – Sep 2013	Bachelor of Science in Psychology, University of Koblenz-Landau, Germany.
Jun 2008	Abitur

Mathias Twardawski

Address: Rohrbacher Straße 168 – 69126 Heidelberg

Phone: +49 176-78779144 – **E-Mail:** mathias.twardawski@gmail.com

PUBLICATIONS

Yudkin, D. A., Rothmund, T., Twardawski, M., Thalla, N., & Van Bavel, J. (2016). Reflexive intergroup bias in third-party punishment. *Journal of Experimental Psychology: General, 145*, 1448-1459.

Twardawski, M., Hilbig, B. E., & Thielmann, I. (in press). Punishment goals in classroom interventions: An attributional approach. *Journal of Experimental Psychology: Applied*.

Twardawski, M., Tang, K. T.-Y., & Hilbig, B. E. (under review). Is it all about retribution? The flexibility of punishment goals.

Twardawski, M., & Hilbig, B. E (under review). Appropriate but unfair? Comparing teachers' and students' perspectives on the treatment of student misbehavior.

Twardawski, M., & Hilbig, B. E (under review). For the greater good? The motivational basis of third-party punishment in children.

CONFERENCE CONTRIBUTIONS

Twardawski, M., Tang, K. T.-Y., & Hilbig, B. E. (2018). Is it all about retribution? On the flexibility of punishment goals. Talk at the 51. Kongress der Deutschen Gesellschaft für Psychologie (DGPs) [51th Congress of the German Psychological Association], Frankfurt, Germany.

Twardawski, M., & Hilbig, B. E. (2018). The discrepancy between students' and teachers' punishment goals: Why students think they are wronged. Symposium Presentation: Proximal processes in classrooms examined from multiple perspectives, MoSAiK Conference: The interplay of theory and practice in teacher education, Koblenz, Germany.

Twardawski, M., Tang, K. T.-Y., & Hilbig, B. E. (2018). Studying laypeople's punishment goals: The neglected role of information salience. Talk at the International Conference on Justice Research (ISJR), Atlanta; Georgia, USA.

Twardawski, M. (2018). Interindividuelle Unterschiede in Bestrafungszielen. Talk at the Doktorandenworkshop der DGPs-Fachgruppe Differentielle Psychologie, Persönlichkeitspsychologie und Psychologische Diagnostik [Meeting for Early-Career Researchers of the German Psychological Association] in Landau, Germany.

Twardawski, M., & Hilbig, B. E. (2017). Die Motivation für Bestrafung: Konkretes Verhalten und mögliche Konsequenzen. Talk at the Moralforschertagung [Meeting of Research on Morality], Leipzig, Germany.

Klein, S. A., Thielmann, I., Henninger, F., Twardawski, M., & Hilbig, B. E., (2017). Public goods versus environment: A social dilemma to investigate the conflict between cooperation and pro-environmental behavior. Poster presented at the 17th International Conference on Social Dilemmas, Sicily, Italy.

Twardawski, M., Hilbig, B. E., Henninger, F., Klein, S. A., & Thielmann, I. (2016). The myth of the little emperor – are only children no more selfish? Poster presented at the 50. Kongress der Deutschen Gesellschaft für Psychologie (DGPs) [50th Congress of the German Psychological Association], Leipzig, Germany.

Twardawski, M., Hilbig, B. E., Henninger, F., Klein, S. A., & Thielmann, I. (2016). Only children, siblings, and the development of selfishness. Poster presented at the 16th Biennal Meeting of the International Society for Justice Research (ISJR) in Canterbury, UK.

Twardawski, M., Yudkin, D. A., Rothmund, T., Thalla, N., Van Bavel, J., J., Bröder, A., & Hilbig, B. (2016). Limited working memory capacity promotes intergroup bias in third-party punishment. In Funke, J., Rummel, J. & Voß, A. (Eds.), Abstracts of the 58th Conference of

Mathias Twardawski

Address: Rohrbacher Straße 168 – 69126 Heidelberg

Phone: +49 176-78779144 – **E-Mail:** mathias.twardawski@gmail.com

Experimental Psychologists. Lengerich: Pabst Science Publishers.

Twardawski, M., & Hilbig, B. (2016). Schulgerechtigkeit als wichtige Einflussquelle des Bildungserfolgs: Ein Vergleich der Bestrafungspräferenzen von Lehrer/innen und Schüler/innen. Talk at the Conference for Early-Career Researchers of the Gesellschaft für empirische Bildungsforschung (GEBF) [Society for Empirical Educational Research] 2016, Berlin, Germany.

Twardawski, M. (2015). People are intuitively cooperative – but because it's morally right, not because it's clever. Talk at the 7th Annual JDM Workshop for Early-Career Researchers, Göttingen, Germany.

Twardawski, M., Yudkin, D. A., Rothmund, T., & Van Bavel, J. (2014). Limited working memory capacity promotes in-group favoritism in punishment. Talk at the 7th Annual JDM Workshop for Early-Career Researchers, Mannheim, Germany.

Twardawski, M., Yudkin, D. A., Rothmund, T., & Van Bavel, J. (2014). Limited working memory capacity promotes in-group favoritism in punishment. In Yudkin, D. A., Van Bavel, J. J., & Rothmund, T., Collective Regulation: Group-level Function of Moral Judgment. Symposium Presentation at the 15th Biennial Conference of the International Society for Justice Research (ISJR), New York City; New York, USA.

WORKSHOPS AND TEACHING

Winter term 2018/2019:	Seminars “Empirical Training II” and “Theoretical and Methodological Basics II”, B.Sc. Psychology, Ludwig-Maximilians-Universität München.
Spring term 2018:	Session on the topic „Descriptive Statistics, Tests und simple Graphics in R“, B.Sc. Psychology, University of Koblenz-Landau.
Winter term 2017/2018:	Workshop on the topic „Reproducible Data Analysis in R – A Tutorial in R Markdown and the “tidyverse” R package”, DGF-Graduate-School „Teaching and Learning Processes“, University of Koblenz-Landau.
	Session on the topic „Preregistration“, B.Sc. Psychology, University of Koblenz-Landau.
Spring term 2017:	Session on the topic „Descriptive Statistics, Tests und simple Graphics in R“, B.Sc. Psychology, University of Koblenz-Landau.
Winter term 2016/2017:	Workshop on the topic „Do’s and Don’ts of Data Analysis – How to avoid (accidental) *p*-hacking“, DGF-Graduate-School „Teaching and Learning Processes“, University of Koblenz-Landau.
	Session on the topic „Descriptive Statistics, Tests und simple Graphics in R“, B.Sc. Psychology, University of Koblenz-Landau.